

Dissertation zur Erlangung des Doktorgrades  
der Fakultät für Chemie und Pharmazie  
der Ludwig-Maximilians-Universität München

**Structural and functional investigation  
of the innate immune receptor MDA5 and its inhibition  
by paramyxoviral V proteins**

Carina Beate Motz

aus

Fulda, Deutschland

2013

### Erklärung

Diese Dissertation wurde im Sinne von § 7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Karl-Peter Hopfner betreut.

### Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, am 14.5.2013

---

Carina Beate Motz

Dissertation eingereicht am	14.5.2013
1. Gutachterin / 1. Gutachter:	Prof. Dr. Karl-Peter Hopfner
2. Gutachterin / 2. Gutachter:	Prof. Dr. Karl-Klaus Conzelmann
Mündliche Prüfung am	8.7.2013

This thesis has been prepared from April 2010 to February 2013 in the laboratory of Prof. Dr. Karl-Peter Hopfner at the Gene Center of the Ludwig-Maximilians-University of Munich (LMU).

**Parts of this thesis have been published:**

**Motz C.**, Schuhmann K.M., Kirchhofer A., Moldt M., Witte G., Conzelmann K.K., Hopfner K.P.  
“*Paramyxovirus V proteins Disrupt the Fold of the RNA Sensor MDA5 to Inhibit Antiviral Signaling.*”  
Science (8 February 2013: Vol. 339, no. 6120, pp. 690–693).

**Parts of this thesis have been presented at an international conference:**

Talk at the 23<sup>rd</sup> Annual Meeting of the Society for Virology, 6<sup>th</sup>–9<sup>th</sup> March 2013, Kiel, Germany.

*„Ein hartnäckiger Begleiter der Erkenntnis ist die  
Unwissenheit über die eigene Unwissenheit.“*

Stanislaw Lem, *Also sprach Golem*

## Table of Contents

<b>1 Introduction.....</b>	<b>1-23</b>
<b>1.1 Pattern recognition receptors are key sentinels of the innate immune system.....</b>	<b>1-5</b>
1.1.1 Universality versus specificity.....	1
1.1.2 Toll-like receptors (TLRs).....	2
1.1.3 NOD-like receptors (NLRs).....	2-3
1.1.4 The AIM2 inflammasome.....	3
1.1.5 Other DNA sensors.....	3
1.1.6 RIG-I like receptors (RLRs).....	3-4
1.1.7 Components of the innate immune system do not act in isolation.....	4-5
<b>1.2 The RIG-I like receptor family senses viruses by recognizing RNA patterns.....</b>	<b>5-10</b>
1.2.1 RLRs share a similar domain organization but function divergently in antiviral immunity.....	5-8
1.2.2 The RD determines the specificity of RLRs for different RNA ligands.....	8-10
<b>1.3 The structural basis for antiviral signal activation by RLRs.....</b>	<b>10-15</b>
1.3.1 RIG-I is activated by a conformational switch that triggers signaling.....	10-12
1.3.2 The activation of MDA5 shows mechanistic similarities and differences to RIG-I.....	12-14
1.3.3 The function of ATP hydrolysis in the activation of RLRs is unknown.....	14-15
<b>1.4 RIG-I and MDA5 use the adaptor MAVS to induce cytokine production.....</b>	<b>15-19</b>
1.4.1 Ubiquitin and the polymerization of MAVS propagate the antiviral signaling cascade.....	15-16
1.4.2 The activated transcription factors NF- $\kappa$ B, IRF3 and IRF7 drive the cytokine production.....	16-17
1.4.3 CARDless LGP2 regulates the antiviral activity of RIG-I and MDA5.....	17-19
<b>1.5 The paramyxoviral V protein is a potent cytokine antagonist.....</b>	<b>19-23</b>
1.5.1 The V protein is a non-structural protein required for viral pathogenesis.....	19-20
1.5.2 The V protein sabotages the innate immune response at several levels.....	20-21
1.5.3 The V protein specifically binds and inhibits MDA5.....	22-23
<b>1.6 Objectives.....</b>	<b>23</b>

<b>2 Material and Methods .....</b>	<b>24-42</b>
<b>2.1 Materials.....</b>	<b>24</b>
<b>2.2 Methods.....</b>	<b>24-42</b>
2.2.1 General molecular biology methods.....	24-25
2.2.1.1 Cloning.....	24
2.2.1.2 Site-directed mutagenesis.....	24
2.2.1.3 Heat shock transformation of <i>Escherichia coli</i> XL1-Blue (K12).....	25
2.2.1.4 Isolation of plasmid DNA from <i>Escherichia coli</i> XL1-Blue (K12).....	25
2.2.1.5 Determination of nucleic acid and protein concentrations.....	25
2.2.2 General bioinformatics methods.....	25
2.2.2.1 Protein sequence alignments and secondary structure predictions.....	25
2.2.2.2 Structure visualization and analysis.....	25
2.2.3 Small scale protein testexpression .....	26
2.2.4 Large scale protein expression.....	27
2.2.5 Large scale purification of MDA5 and PIV5 V proteins and their complexes.....	28-29
2.2.6 ATPase assays.....	29
2.2.7 Static light scattering (SLS) measurements.....	30-31
2.2.8 Small-angle X-ray scattering (SAXS) of protein solutions .....	31
2.2.9 Reductive methylation .....	32
2.2.10 Large scale purification of Fab fragments and triple complex formation .....	32
2.2.11 Crystallization of the ssMDA5:PIV5 V protein core complex.....	32-33
2.2.12 Structure determination.....	33-41
2.2.12.1 Principles of X-ray structure analysis.....	33-34
2.2.12.2 Principles of SAD phasing.....	35-40
2.2.12.3 Structure determination of the ssMDA5:PIV5 V protein core complex.....	40-41
2.2.13 Analytical size exclusion chromatography.....	41
2.2.14 Differential scanning fluorimetry experiments .....	41
2.2.15 Electrophoretic mobility shift assays (EMSAs).....	41-42
2.2.16 Electron microscopy .....	42

<b>3</b>	<b>Results.....</b>	<b>43-65</b>
3.1	Recombinantly expressed full length MDA5 is a functional protein.....	43-45
3.2	Optimization of protein constructs for crystallization.....	46-49
3.3	Crystallization of the ssMDA5:PIV5 V protein core complex.....	49-52
3.4	The paramyxoviral V protein disrupts the fold of the MDA5 SF2 ATPase domain.....	52-56
3.5	Differences in residue conservation enable RIG-I to evade V protein recognition.....	57-59
3.6	The PIV5 V protein has to unfold itself to disrupt the fold of MDA5.....	59-62
3.7	V protein binding has dual impact on the function of MDA5.....	63-65
<b>4</b>	<b>Discussion.....</b>	<b>66-80</b>
4.1	A structure-derived model for MDA5 inhibition by paramyxoviral V proteins.....	66-72
4.1.1	Which energy source drives the unfolding processes that accompany V protein binding?.....	69-70
4.1.2	Known $\beta$ -strand displacement mechanisms in other biological processes.....	70-72
4.2	Is the RD of RLRs able to switch between two binding modes?.....	72-73
4.3	V protein binding to other RLR family members and its functional impact.....	73-78
4.3.1	Engineering of RIG-I inhibition by paramyxoviral V proteins.....	75-77
4.3.2	Indirect inhibition by paramyxoviral V proteins via LGP2.....	78
4.3.3	Viral pressure on the evolution of RLRs.....	78
4.4	The paramyxoviral V protein as a universal immune targeting tool.....	79-80
<b>5</b>	<b>Summary.....</b>	<b>81-82</b>
<b>6</b>	<b>References.....</b>	<b>83-94</b>
<b>7</b>	<b>List of abbreviations.....</b>	<b>95-97</b>
<b>8</b>	<b>Acknowledgments.....</b>	<b>98</b>
<b>9</b>	<b>Curriculum vitae.....</b>	<b>99</b>

## 1. Introduction

### 1.1 Pattern recognition receptors are key sentinels of the innate immune system

Any organism that passes the physical and chemical barriers of the body encounters a further level of defense, the innate immune system of the host. Its crucial role in the early recognition of invading pathogens as bacteria, viruses, protozoa and fungi is emphasized by the fact that defects affecting the components of the innate immune response are almost always lethal (1). The selective pressure imposed on multicellular organisms by infectious microorganisms directed the evolution of a large class of germline-encoded receptors, known as pattern recognition receptors (PRRs). This class of innate immune receptors, which comprises membrane-bound as well as soluble receptor proteins, was named after their capability to sense particular molecular signatures on pathogens called pathogen-associated molecular patterns (PAMPs). In this capacity, PRRs represent an essential key component of the host innate immune defense (2, 3).

#### 1.1.1 Universality versus specificity

The term “pattern”, which Janeway introduced already in 1989, reflects the basic molecular concept of PAMP recognition by host PRRs. PAMPs are characterized by a certain degree of structural invariance since they represent molecular structures that are essential for the survival of the pathogen. The resulting evolutionary conservation of these structures comprises that they are shared by large groups of pathogens. For example, lipopolysaccharides (LPS) can be found in the outer membrane of all gram-negative bacteria and share the conserved lipid A. Serving as common molecular pattern for the PRR TLR 4, lipid A enables the host to detect the presence of any gram-negative bacterium. In this way, a relatively small number of germline-encoded receptors can recognize a broad range of molecular structures associated with pathogens (3).

However, the interaction between conserved PAMPs and PRR is not only characterized by universality, but also by specificity. To eliminate the invading pathogen and to induce an antipathogenic state of the host, the innate immune system employs various defense mechanisms that involve e.g. the production of proinflammatory cytokines, activation of the complement system, induction of apoptosis, phagocytosis or autophagy (2). To avoid self-damage of the host cells and tissues, the patterns evolutionary selected for the recognition by PRRs must have been absolutely distinct from the self-antigens. This selected component of specificity allows PRRs to distinguish infectious nonself from noninfectious self and thereby contributes to the required and tight regulation of the immune response (3).



### 1.1.2 Toll-like receptors (TLRs)

The discovery of PRRs is linked to the identification of the Toll protein in *Drosophila* species in 1985. The identification of mammalian Toll-like receptors (TLRs) succeeded in the mid 1990s based on homology to the *Drosophila* protein. The finding that they mediate the recognition of pathogens by the innate immune system exposed the existence of PRRs. The TLRs are to date the most extensively studied PRRs. A total of 10 TLR family members have been identified in humans and 12 in mice. They all belong to the type I transmembrane glycoproteins and bind PAMPs via leucine rich repeats (LRRs) in their extracellular portion. A cytoplasmic Toll/interleukin-1 receptor (TIR) domain provides downstream signaling. Based on the nature of the recognized PAMPs, the TLRs can be divided into subfamilies: TLR1, 2, 4 and 6, which are primarily expressed on the cell surface recognize lipids derived from bacteria and viruses. In contrast, TLR3, 7, 8 and 9 are exclusively located within endocytic compartments and bind to bacterial and viral nucleic acids (2, 4, 5).

### 1.1.3 NOD-like receptors (NLRs)

The recognition by membrane-bound TLRs is restricted to extracellular pathogens. Since the viral replication cycle for instance proceeds inside the cell, TLR recognition is supported by various cytoplasmic receptors that contribute in an essential way to an effective immune response (6). The family of NOD-like receptors (at least 23 members in humans) mediates sensing of a broad ligand spectrum from all classes of pathogens within the cytoplasm of cells. These receptors consist of three domains: The N-terminus carries a signaling domain as for example a pyrin domain (PYD) or a caspase recruitment domain (CARD). These domains are known to function as protein interaction modules and therefore thought to direct homo- or heterotypic binding. The signaling domain is followed by a nucleotide-binding and oligomerization domain (NOD) that is believed to trigger an ATP-dependent self-oligomerization of the receptor molecule upon ligand binding. The C-terminus accommodates a series of LRRs which are, as in the case of TLRs, implicated in the binding of the PAMP (4, 7).

NOD1 and NOD2 were the first described NLRs that are involved in cytoplasmic microbial sensing. Both molecules are activated by peptidoglycan components of the bacterial cell wall and trigger NF- $\kappa$ B-dependent production of inflammatory cytokines (7). Other stimuli are able to drive the assembly of certain NLRs (e.g. NALP, NALP3, NLRC4, NAIP5) into large, multiprotein signaling complexes, named “inflammasomes”(4). The formation of an inflammasome complex involves the scaffold protein ASC (apoptosis-associated speck-like protein containing a CARD). Its N-terminal PYD and its C-terminal CARD enables ASC to function as a bipartite adaptor molecule that recruits members of the NLR family as well as non-NLR proteins (7). Among ASC recruited non-NLR proteins are various

procaspases that undergo proteolytic cleavage and thus activation upon inflammasome association. The in- complex activated caspases in turn cleave substrates to produce mature inflammatory cytokines (the substrates include for example pro-IL-1 $\beta$  and pro-IL-18) or to initiate apoptosis (4).

To date it is still unresolved whether NLRs activation is induced by direct binding of the PAMP ligand or rather by indirect sensing of particular stress-activated signaling pathways that go along with invasion of cells by pathogens. In this context, it has been proposed that the stress events potassium efflux and generation of reactive oxygen species (ROS) affect the NLR response, but the precise activation mechanism of NLRs still remains unclear (5, 6).

### 1.1.4 The AIM2 inflammasome

Although AIM2 shares the characteristic N-terminal pyrin signaling domain with some NLRs, AIM2 (absent in melanoma 2) does not represent a typical NLR. Unlike NLRs, AIM2 harbors a C-terminal HIN-200 domain that enables the receptor molecule to bind directly to viral dsDNA within the cytoplasm. The PYD allows activated AIM2 the formation of an inflammasome via a homotypic interaction with the adaptor protein ASC. The AIM2 inflammasome is associated with procaspase-1 whose activated form caspase-1 converts pro-IL-1 $\beta$  and pro-IL-18 into mature cytokines (4-6, 8).

### 1.1.5 Other DNA sensors

It is known that cytoplasmic DNA can induce the production of proinflammatory cytokines independent from TLRs or AIM2, pointing towards the existence of additional DNA sensors within the cell. The protein STING (stimulator of interferon genes) localized at the ER-membrane is appointed to be a critical component in the signaling pathway of dsDNA recognition. Since STING does not show a direct binding of dsRNA, it is more likely that it functions downstream as a signaling adaptor protein of an yet unidentified DNA sensor (4, 5, 9). The PYHIN (pyrin and HIN domain-containing protein) protein family member IFI16 (interferon-inducible protein 16) might be a promising bridging factor between dsRNA recognition and STING signaling. It contains two DNA-binding HIN domains and has been shown to be recruited to synthetic dsDNA and STING in the cytoplasmic compartment (8).

### 1.1.6 RIG-I like receptors (RLRs)

The family of RIG-I like receptors comprises the three members RIG-I (retinoic acid inducible gene I), MDA5 (melanoma differentiation associated protein 5) and LGP2 (laboratory of genetics and physiology 2). This family plays a crucial role in the detection of viral RNAs in the cytoplasm of infected cells. Their

function as viral RNA sensors, which is explained in detail in sections 1.2, 1.3 and 1.4 of the introduction, is supposed to be assisted by the OAS/RNaseL pathway. In the presence of intracellular viral dsRNA, OAS (2'-5' oligoadenylate synthase) is activated and produces a specific nucleotide second messenger, namely 2'-5' oligoadenylate. In response to this second messenger, RNaseL cleaves and processes viral nonself as well as cellular self RNAs. The RNaseL products are believed to serve as potent RNA ligands for members of the RLR family (1, 10, 11).

Beside their function as RNA receptors, RLRs are also able to act as indirect DNA sensors by recognizing viral transcripts produced by the RNA polymerase III (Pol III). Pol III normally transcribes cellular DNA to synthesize ribosomal 5S rRNA, pre-tRNA and other small RNAs. In infected cells, viruses employ the host Pol III to generate small, noncoding RNAs that can antagonize the host immune response, but can be also recognized by RLRs. By transcribing a particular region of the DNA virus genome, Pol III synthesizes 5'-triphosphate ends, which are known to effectively activate RIG-I (see section 1.2). Thus, the Pol III-RIG-I pathway nicely illustrates the evolutionary arms race between the host and the virus (8).

### 1.1.7 Components of the innate immune system do not act in isolation

Studies suggest that the introduced PRR families are not sufficient to explain the protective immune response that can be observed against a wide range of various infectious agents. Indeed, the numbers of PRRs, but also their ligands continue to increase, indicating that many important PRRs are still undiscovered (4, 7).

But so far signaling pathways triggered by PRRs seem to share a general concept: Signaling domains within the receptor molecule allow the interaction with certain signaling adaptor proteins (e.g. Myd88, Trif, TRAM, Mal, MAVS, RIP2), which operate in cascades and activate terminal effector proteins such as caspases (e.g. caspase-1 or caspase-5) and transcription factors (e.g. NF- $\kappa$ B, IRF3 and IRF7). The effector proteins induce finally the production of type I interferons and other various proinflammatory cytokines or chemokines as well as anti-viral proteins. All these signaling pathways do not act in isolation, but have been shown to be synergistic and cooperative. This interplay between the different PRR families seems to permit the effective orchestration of the innate immune response (7).

Not only the components of the innate immunity are interlinked, but there is also intensive communication between the pathways of the innate and adaptive immune system. Besides its role in the early recognition of pathogens, the innate immune system contributes fundamentally to the proper activation, maturation, priming and shaping of the adaptive immune response. In this capacity, the innate immune system is not longer regarded as a primitive evolutionary rudiment that functions in separation of

the adaptive immune system. Instead, the innate immunity and the adaptive immunity are perceived as two tightly connected systems that act in concert to provide efficient and broad protection of the host organism (1, 2).

### **1.2 The RIG-I like receptor family senses viruses by recognizing RNA patterns**

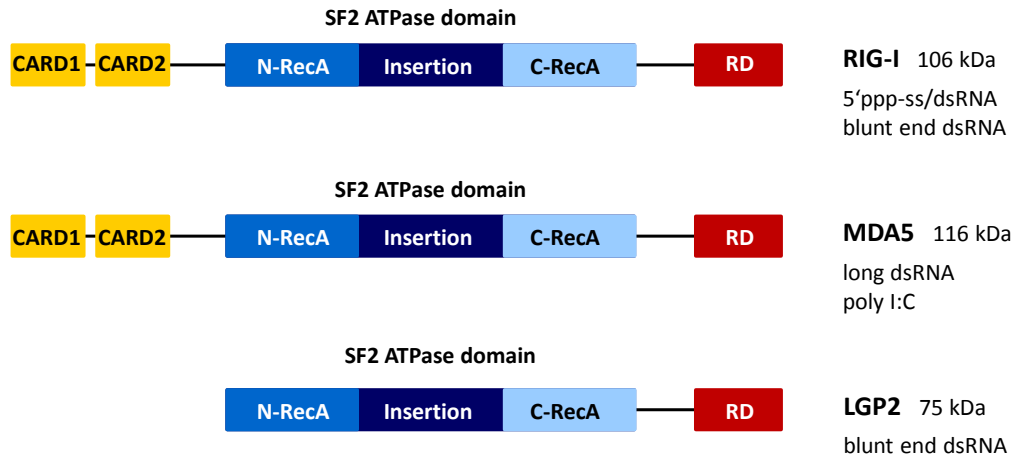
In addition to TLRs, the RIG-I like receptor (RLR) family represents the second class of PRRs that senses RNA PAMPs and plays a major role in inducing an antiviral response. But unlike the membrane-bound TLRs, the three RLR family members RIG-I, MDA5 and LGP2 are localized and distributed in the cytoplasm where they bind to specific RNA patterns that occur in the host cell during a virus infection (see also sections 1.1.6 and 1.2.1). RIG-I, the founding member and eponym of the family, as well as MDA5 were initially characterized as RNA helicases that showed ATPase activity upon binding to dsRNA and were therefore assumed to be involved in the induction or mediation of an innate antiviral response (12, 13). Various studies have since revealed RIG-I and MDA5 as major factors in virus recognition and activation of antiviral signaling pathways (14, 15). LGP2, which was originally identified as a gene of the Stat3/5 locus highly expressed in mammary tissues (16), complements the RLR family and is believed to play an important role in the modulation of the signaling by its homologs RIG-I and MDA5 (17-23) (see also section 1.4.3).

#### **1.2.1 RLRs share a similar domain organization but function divergently in antiviral immunity**

RIG-I, MDA5 and LGP2 belong to the superfamily 2 (SF2) helicases or ATPases. They share a central DECH box RNA helicase domain (named after the conserved amino-acid sequence of the Walker B motif) that exhibits ATP hydrolysis as well as RNA binding capacity (14, 24, 25). The DECH box RNA helicase domain contains the two characteristic RecA-like domains (also known as Rossmann fold), that define the SF1 and SF2 helicases (see figure 1). These two domains form an active site for ATP binding and hydrolysis at their interface by clustering conserved motifs in a central core region. DECH box helicases possess eight conserved motifs (namely motifs I (Walker A motif), Ia, Ib, II (Walker B motif), III, IV, V and VI) that were shown to be involved in the regulation of the ATPase activity and some of the motifs also mediate RNA binding (24, 26).

RIG-I like helicases gain uniqueness and functional specificity through their accessory domains that complement the helicase core (see figure 1). For instance, the two RecA-like domains are separated by an unusual  $\alpha$ -helical insertion domain that resembles the archaeal Hef DNA helicase (27). In

addition, the second RecA-like domain is followed by a C-terminal extension that contains a zinc-binding domain with structural homology to GDP/GTP exchange factors of Rab-like GTPases (28). This so called regulatory domain (RD) as well as the alpha-helical insertion domain have been shown to participate in the binding of RNA ligands in an essential way (26, 29-31) (see also sections 1.2.2 and 1.3). However, sole RNA binding by RLRs is not sufficient to induce an antiviral activity, but relies on the presence of a CARD interaction motif. RIG-I and MDA5 (but not LGP2) are equipped with two tandem CARDS, which are localized at the N-terminus and which serve as signaling domains that mediate binding to the adaptor protein MAVS (see also section 1.3 and 1.4). Their crucial role in triggering antiviral response is emphasized by the fact that overexpression of the isolated RIG-I CARDS is sufficient to confer IFN production, showing that the CARDS are indeed responsible for antiviral signaling (12).



**Figure 1:** Domain organization of the three RIG-I like receptor family members RIG-I, MDA5 and LGP2 with indicated molecular weight and specified RNA ligands. CARD = caspase recruitment domain; SF2 = helicase superfamily 2, RD = regulatory domain or RNA binding domain. Adapted from (14, 25).

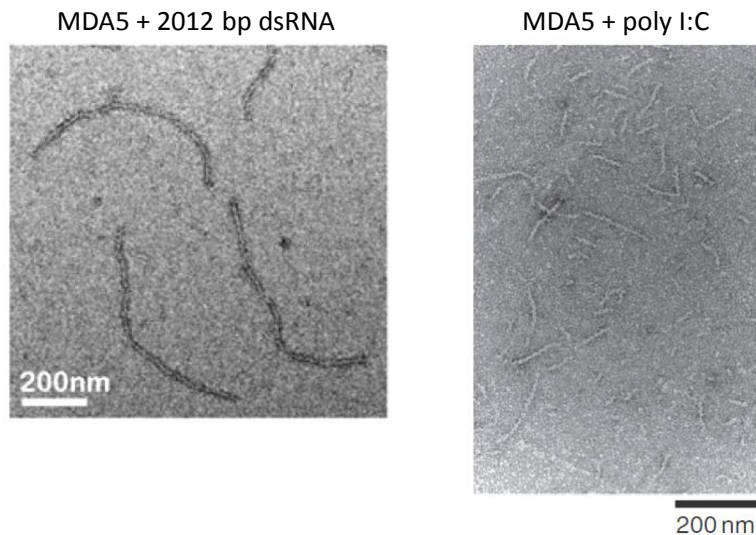
Despite the similarities that they share in sequence and domain organization, RLRs show divergent function concerning the groups of viruses that they recognize. RIG-I mainly detects negative-stranded viruses (such as influenza virus (*Orthomyxoviridae*), Sendai virus (*Paramyxoviridae*), Newcastle disease virus (*Paramyxoviridae*) and vesicular stomatitis virus (*Rhabdoviridae*)), but also some positive-stranded viruses (such as hepatitis C virus (*Flaviviridae*) and Japanese encephalitis virus (*Flaviviridae*)). In contrast, MDA5 responds primarily to double-stranded and positive-stranded viruses, especially to members of the *Picornaviridae* family (such as encephalomyocarditis virus (EMCV), Theiler's murine

encephalomyelitis virus (TMEV), coxsackievirus B3 (CVB3)) and to murine norovirus (*Caliciviridae*) (14, 32-38).

This divergence in virus recognition by RIG-I and MDA5 is reflected in the actual structural RNA patterns, which RLRs detect on the molecular level and which allow them to distinguish different RNA viruses. In 2006, two independent studies identified 5'-triphosphate (5'-ppp) containing RNAs as RIG-I ligands (39, 40). Indeed, many RNA virus that are recognized by RIG-I carry genomes that contain 5'-ppp ends or synthesize RNAs with 5'-ppp ends during their replication cycle in the host cell. The particles of negative-stranded viruses import a RNA-dependent RNA polymerase (RdRp) that allows those viruses the immediate transcription of their genome after entering the cell. Most RdRps do not rely on a primer to initiate synthesis, resulting in transcripts that exhibit a triphosphate group at their 5'-end (41, 42). But activation of RIG-I does not necessarily require viral replication within the cell. It has been shown that RIG-I is able to induce IFN response already before viral RNA synthesis has started by detecting incoming, encapsidated RNA virus genomes, which carry a 5'-ppp dsRNA "panhandle" structure (43, 44).

The 5'-ppp moieties are specific for viral RNA genomes or transcripts as most of the 5'-ppp groups that emerge from host transcription are either removed by adding a cap structure (in case of mRNA) or by processing (in case of tRNA and rRNA). To that effect, 5'-ppp moieties allow RIG-I to distinguish between self and non-self (34). Albeit RIG-I shows highest affinity for 5'-ppp dsRNA, it is known that RIG-I can also recognize and bind blunt-ended dsRNA (as LGP2 does, see also sections 1.2.2 and 1.4.3) (31, 45-47). This indicates that other, not yet identified and maybe multiple features of RNA, which differ between viral and cellular RNAs, might be involved in the discrimination of self and non-self by RLRs (48).

Little is known about the nature of the RNA patterns that activate MDA5. Its physiological ligands are scarcely characterized and as yet poorly defined. Uncapped 5'-ppp RNA is absent in viruses known to be detected by MDA5. Rather, it has been suggested, that ribose 2'-O-methylation provides a molecular signature for MDA5 to differentiate between self and non-self mRNA (49). However, it is generally assumed that MDA5 is activated by long, double-stranded RNA molecules, which are regarded as "intrinsically" non-self (47). It is known for a long time, that MDA5 can be activated by transfection of poly(I:C), a synthetic dsRNA analog exceeding 2 kbp in length. RNase III digestion of poly(I:C) to ~300 bp fragments converts poly(I:C) from a MDA5 activator to a RIG-I activator, showing that RIG-I does not necessarily rely on the presence of 5'-ppp groups and prefers shorter RNA ligands as MDA5 (35, 47). This original finding, that MDA5 recognizes long dsRNA, is supported by several further studies. Additional transfection-based experiments showed that MDA5 responds to the L segments of the



**Figure 2:** Filament formation of MDA5 on long dsRNAs visualized by negative-stain electron microscopy (53, 54).

reovirus genome (4 kbp in length) with induction of type I interferon production (50). Furthermore, the very large RNA species generated during EMCV or vaccinia virus infection were identified as specific MDA5 activators, indicating that MDA5 requires higher-order RNA structures to function (36). Two other studies proposed that MDA5 detects the double-stranded replicative form (7.5 kbp in length) produced in cells infected with picornaviruses or other positive-stranded RNA viruses (51, 52). Finally, two independent studies reported that MDA5 assembles into long filaments along dsRNA that can be visualized by negative-stain electron microscopy (see figure 2). These studies also provided evidence that the observed filament formation relies on a cooperative binding mode of MDA5. The ability for cooperative filament assembly seems to be a specific hallmark of MDA5 as RIG-I is in the position to tightly bind dsRNA as a monomer and shows little or no cooperativity in dsRNA binding (53, 54). This illustrates once more the distinct but complementary role of RIG-I and MAD5 in virus recognition.

### 1.2.2 The RD determines the specificity of RLRs for different RNA ligands

It is known that the SF2 ATPase domain and the RD act in concert to permit proper RNA binding and recognition by RLRs (26, 29, 30). However, the RD is the actual domain that mediates high affinity binding of RLRs to their RNA ligands and also specifies this binding. Its predominant role in binding is expressed in the observation that the ATPase domain itself is not able to bind RNA under physiological conditions, whereas the RD shows high affinity for RNA also in isolation (31, 55-57). The RIG-I RD was firstly identified in limited proteolysis experiments and found to exhibit high affinity for 5'-ppp dsRNA

as well as blunt-ended dsRNA determining the RD as actual sensor for these ligands (28). The subsequent crystal structures of the RDs of all three RLR members in complex with their RNA ligands revealed how binding of and specificity for different RNA patterns is achieved at the atomic level (see figure 3).

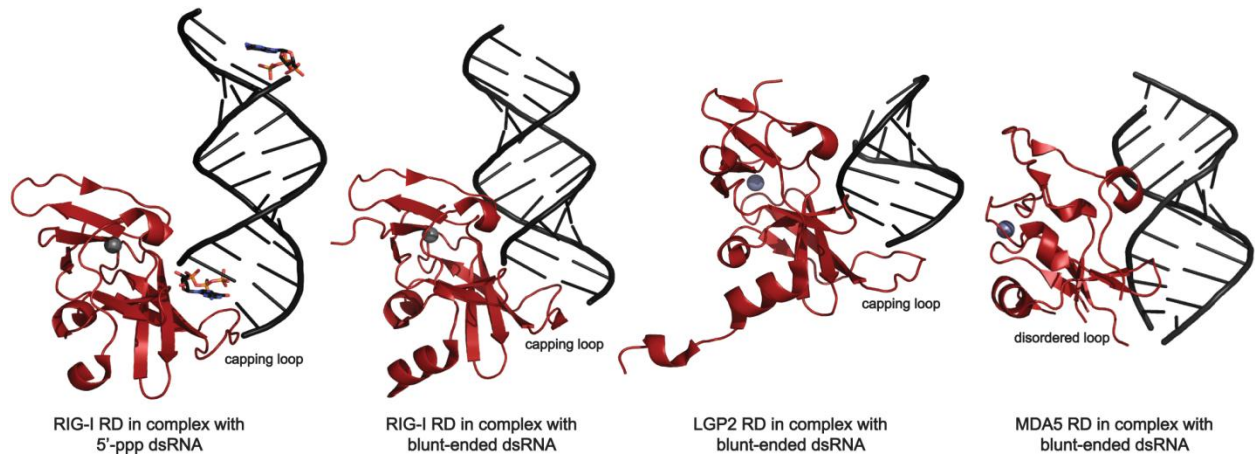
The structure of the RIG-I RD in complex with 5'-ppp dsRNA showed how a loop rigidly caps the end of the dsRNA using a stacked phenylalanine while a lysine-rich basic cleft within the RIG-I RD sequesters the 5'-ppp group (58). The structure of the RIG-I RD in complex with blunt-ended dsRNA revealed that the RD binds triphosphorylated and non-phosphorylated dsRNA in significantly different orientations using a distinct but overlapping set of residues. Specific interactions between the 5'-ppp moiety and the RIG-I RD define how the RD binds 5'-ppp dsRNA, while interactions between the phosphodiester backbone and the RD determine how it binds blunt-ended dsRNA. In comparison to the 5'-ppp dsRNA, the blunt-ended dsRNA swings towards the RIG-I RD in order to increase its contacts with the protein, whereas the conformation of the RD itself is almost identical in both complexes. This indicates that only slight conformational adjustments of the RIG-I RD specify the binding and recognition of the two RNA forms (46).

In case of the LGP2 RD, the blunt-ended dsRNA swings even more towards the RD and also rotates along its helical axis relative to the position of the blunt-ended dsRNA bound to the RIG-I RD. This different orientation results in a further extension of the interaction interface between the RNA and the LGP2 RD. In addition, the basic 5'-ppp binding cleft within the RIG-I RD becomes hydrophobic for LGP2 and also the loop that caps the exposed base pairs shows a slight different conformation. These structural distinctions determine altogether the different binding mode of the RIG-I RD and the LGP2 RD (31, 46, 58).

To date, a structure of the isolated MDA5 RD in complex with its RNA ligand is not available, indicating that its RNA binding capability relies to a greater degree on the presence of the SF2 ATPase domain than in cases of the RIG-I RD and LGP2 RD (see also section 1.3.2) (59). Indeed, the MDA5 RD shows the lowest RNA affinity among all three homologs (59) and crystallization of a truncated MDA5 version containing the RD and the SF2 ATPase domain in complex with dsRNA succeeded. This structure revealed that the MDA5 RD recognizes the stem region of the dsRNA ligand rather than its ends. The loop, that the RIG-I RD uses to rigidly cap the end of the dsRNA, is disordered in the case of MDA5. Instead, residues on the flat surface of the MDA5 RD mediate dsRNA stem recognition. In this way, the MDA5 RD adds a decisive contribution to the capability of MDA5 molecules to assemble into filaments along dsRNA (see also section 1.3.2) (30). The comparison of the MDA5 RD and the RIG-I RD further reveals that the basic 5'-ppp binding cleft within the RIG-I RD becomes less basic and shows a outwards protrusion in MDA5, explaining why MDA5 is not able to recognize 5'-ppp moieties (58).



In this sense, the RD routes the selectivity of the RLR family members for different RNA structures and therefore permits the detection of largely distinct groups of viruses and viral RNAs.



**Figure 3:** Crystal structures of the RDs of all three RLR family members in complex with their RNA ligands. The RD is shown in red, the structural zinc ion in grey and the RNA ligand in black. Adapted from (30, 31, 46, 58).

### 1.3 The structural basis for antiviral signal activation by RLRs

A series of recent structural studies on RIG-I and MDA5 in isolation and in complex with dsRNA gave important mechanistic insight into the function of these RLRs. The crystal structures presented in these studies expand the knowledge that was gained from the structural work on the isolated RLR RDs (see section 1.2.2) and provide a more complete picture regarding RNA recognition, activation, regulation and antiviral signaling induction by RLRs. These studies also highlight structural differences between the two RLR family members RIG-I and MDA5 that contribute to their divergent function (see section 1.2.1).

#### 1.3.1 RIG-I is activated by a conformational switch that triggers signaling

RIG-I is one of the first SF2 helicases that were cocrystallized with an RNA substrate. The structural comparison of the RNA-bound form of RIG-I with the ligand-free state revealed that RIG-I is activated by a conformational switch induced by the binding of the RNA (see figure 4). In absence of an RNA ligand, RIG-I shows a rather open and elongated conformation. The RD is supposed to be flexible and to move freely on the linker that connects the RD with the C-terminal RecA-like fold of the SF2 ATPase

domain. In contrast, the CARDS are forming a rigid unit with the  $\alpha$ -helical insertion of the ATPase core. Packing against the insertion region, the CARDS are only partially accessible and simultaneously block the RNA binding site of the insertion. In this manner, the CARDS ensure that RIG-I is kept in an inactive state in the absence of RNA species indicating a viral infection (60).

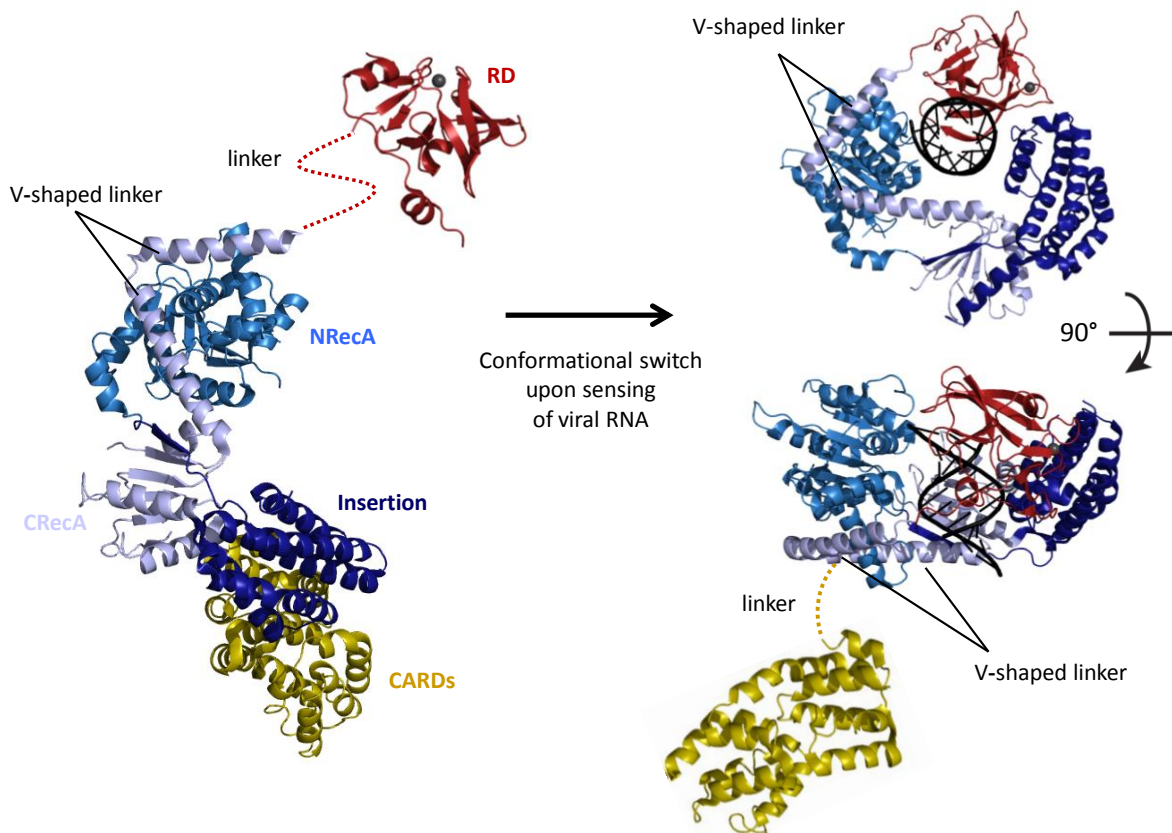
Upon binding of the RNA substrate, the RIG-I molecule undergoes a distinct conformational change accompanied by fundamental rearrangements of the RIG-I domains (see figure 4). The RD and the SF2 ATPase domain act in concert to fully encircle the helix of the dsRNA ligand. The two RecA-like folds come together to create a mature dsRNA- and active ATP-binding site. The RD is not longer flexible, but captured in a compact ring structure, whereas the CARDS are displaced from the insertion region by the incoming RNA ligand. They are now fully accessible and supposed to freely move on the long linker situated between the CARDS and the N-terminal RecA-like fold of the SF2 ATPase domain (26, 29, 60).

The process of binding RNA appears to be highly complex as multiple sites within the RIG-I molecule are involved. In fact, all three subdomains of the ATPase core contribute to the binding of the RNA substrate. In particular, residues of the conserved motifs Ia and Ib within the N-terminal RecA-like fold undergo contacts with the RNA, whereas the C-terminal RecA-like fold contributes with residues of the motifs IV and V. These and other specific contacts are restricted to the phosphodiester backbone of the dsRNA and allow no sequence-specific recognition (26, 29).

The RD participates in this network of backbone interactions, but additionally caps the end of the dsRNA as previously observed (see section 1.2.2). It is supposed that the RD functions as a capture domain whose high affinity for RNA allows initial binding of RNA ligands. This initial binding is thought to increase the local concentration of dsRNA in proximity of the SF2 ATPase domain and thereby to facilitate the formation of the closed and active RNA clamp structure. The conformational switch, that is associated with this step, forces finally the release of the tandem CARDS leading to a full activation of RIG-I (see also section 1.4) (60).

The RIG-I molecule is furthermore characterized by a striking V-shaped linker that is composed of two long  $\alpha$ -helices. This linker bridges the C-terminal RecA-like fold and the RD and simultaneously grips an  $\alpha$ -helix that projects from the N-terminal RecA-like fold. Consequently, the V-shaped linker establishes extensive interactions and a mechanical connection between these three domains in the closed ring formation (see figure 4). The linker is for that reason assumed to facilitate the communication between the domains within the RNA clamp structure. It likely acts as a mechanical transmitter that coordinates and adjusts the relative orientations of the C- and N-terminal RecA-like fold and the RD during the step of RNA binding. In this way, the V-shaped linker does not only contribute to the synergy

between the RD and the SF2 ATPase domain concerning RNA binding, but also assists in the release of the CARDS that represents the actual and final trigger for antiviral signaling (see also section 1.4) (26, 29, 60).



**Figure 4:** Crystal structures of RIG-I in the inactive, ligand-free state (left) and in the active, ligand-bound form (right). The positions of the flexible linked RD and CARDS are just illustrated. The RD is shown in red, the CARDS in yellow and the three subdomains/folds of the SF2 ATPase domain are shaded in blue. The structural zinc ion is highlighted in grey and the dsRNA in black. Adapted from (26, 29, 60).

### 1.3.2 The activation of MDA5 shows mechanistic similarities and differences to RIG-I

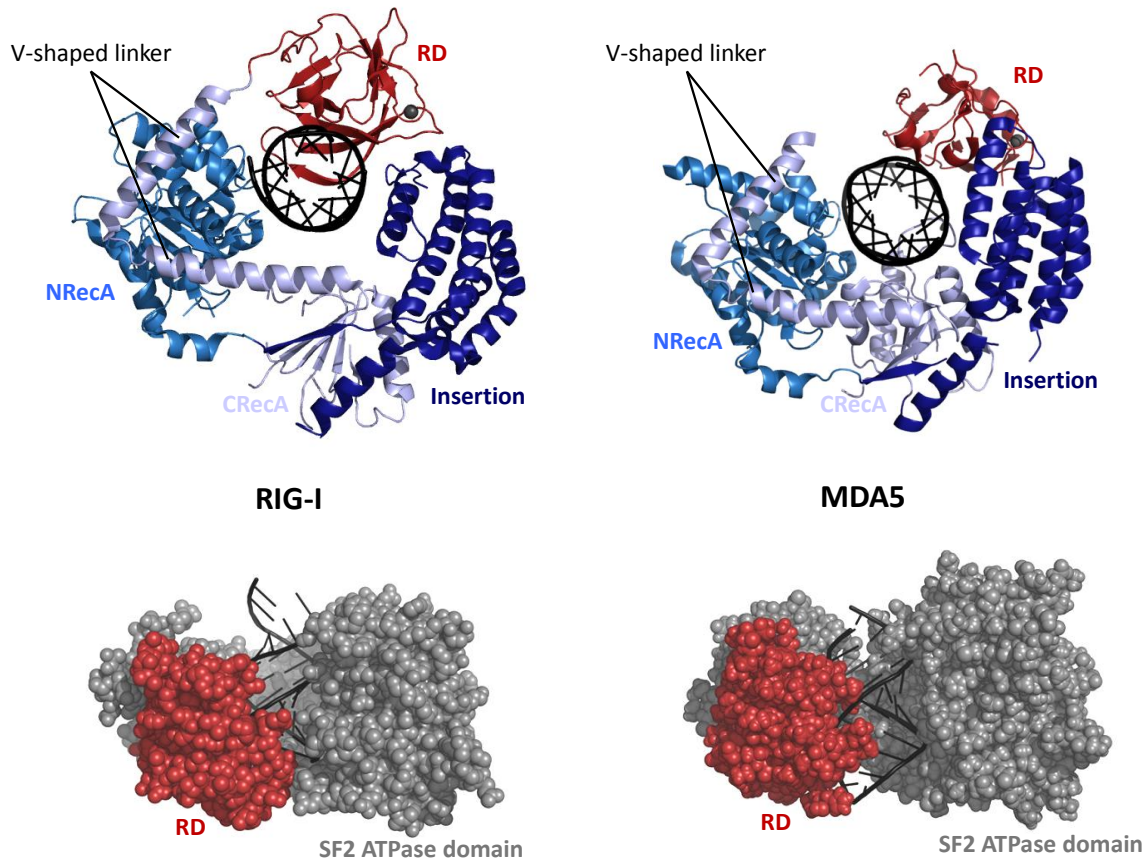
In 2012, the crystallization of the MDA5 RNA clamp structure succeeded. This structure shows high similarity to the corresponding RIG-I RNA clamp structure concerning the overall architecture and their arrangement of the RD and the three subdomains of the SF2 ATPase domain (see figure 5). The RD, N- and C-terminal RecA-like folds, as well as the alpha-helical insertion form a closed ring that completely surrounds the dsRNA as already observed for RIG-I. All four units participate in the binding of the RNA ligand and again the contacts are limited to the phosphodiester backbone and do not allow a sequence-

specific RNA recognition. The MDA5 and RIG-I RNA clamp structure also share the V-shaped linker between the RD and the C-terminal RecA-like fold that is assumed to be important for relaying information between the sensory RD and the catalytic/mechanical devices of the ring (N- and C-terminal RecA-like folds) (see section 1.3.1). All these similarities indicate that related mechanisms underlie the activation of MDA5 and RIG-I (30).

However, the MDA5 RNA clamp structure exhibits subtle differences relative to the RIG-I structure. These structural distinctions have significant impact on the behavior of MDA5 on the molecular level and determine its function. As already emphasized in section 1.2.2, the MDA5 RD recognizes the stem of the dsRNA ligand rather than its ends and this binding mode is influenced by the SF2 ATPase domain of MDA5. In detail, a specific  $\alpha$ -helix within the insertion region positioned at the interface with the RD significantly differs in length between RIG-I and MDA5. This length difference is conserved indicating its functional importance. Within the RNA clamp structure, the longer  $\alpha$ -helix of the RIG-I insertion pushes the RIG-I RD towards the bottom tip of the N-terminal RecA-like fold, which results in an asymmetrically closed O-shaped ring around dsRNA (see figure 5). In contrast, the shorter  $\alpha$ -helix within the MDA5 insertion allows this subdomain to provide a docking site for the MDA5 RD to be positioned near the dsRNA stem. As a consequence, MDA5 forms a more open C-shaped ring around dsRNA with a gap between the MDA5 RD and the N-terminal RecA-like fold (see figure 5). Thus, the MDA5 insertion controls the precise positioning of the MDA5 RD for efficient recognition of the dsRNA stem. The stem is additionally detected by the C-terminal RecA-like fold that inserts a conserved loop into the major groove. These indirect and direct contributions of the MDA5 SF2 ATPase domain to dsRNA stem recognition provide the atomic basis that allows MDA5 molecules to cooperatively assemble into filaments along dsRNA (30) (see also section 1.2.1).

The monomeric MDA5 RNA clamp structure gives no information concerning the MDA5:MDA5 interface and the intermolecular contacts established during filament formation on dsRNA. First structural insights into the cooperative assembly of MDA5 along dsRNA and were obtained by three-dimensional reconstruction of the filamentous MDA5 architecture employing electron microscopy. These studies revealed that MDA5 assembles in a head-to-tail fashion by forming extensive intermolecular contacts between parts of the SF2 ATPase domain, which in particular involve the N- and C-terminal RecA-like folds. Beyond that, the structure reconstruction of the MDA5 filament proposes a key role for the MDA5 RD in cooperative assembly. The data suggest that the RD of one MDA5 ring invades an adjacent MDA5 ring by contacting its insertion as well as its N-terminal RecA-like fold. Suchlike interactions would enable MDA5 to assemble in intertwined rings along dsRNA and hence to generate stable filaments. The CARDs are fully accessible within this arrangement and are supposed to form a regular array at the

outside of the filament. This fibril form is believed to represent the fully activated state of MDA5 that allows efficient antiviral signaling (see section 1.4) (30, 61).



**Figure 5:** Comparison of the RIG-I and MDA5 RNA clamp structure. The RD is shown in red, the three subdomains/folds of the SF2 ATPase domain are shaded in blue (upper panels) or are shown in grey (lower panels). The structural zinc ion is highlighted in grey and the dsRNA in black. Adapted from (29, 30).

### 1.3.3 The function of ATP hydrolysis in the activation of RLRs is unknown

Despite ongoing progress in the functional understanding of RLRs, the precise role of ATP hydrolysis in RLR mediated antiviral signaling remains a major question. It is known that sole RNA binding by RLRs is insufficient to trigger signaling. MDA5 mutants are affected in their capability to hydrolyze ATP but still able to form filaments show a clearly reduced signaling activity in living cells (30). It is further known that simultaneous binding of ATP increases the affinity of RLRs for dsRNA, but the actual step that the hydrolysis of ATP drives in the course of RLR activation remains elusive (60). In contrast to bona fide helicases that use ATP to unwind nucleic acid duplexes, RIG-I is rather assumed to translocate along

dsRNA. This translocation activity of the receptor is supposed to contribute to antiviral signal generation by inducing further changes in protein conformation (62). In this sense, RIG-I and MDA5 might consume ATP to adopt a conformational state that allows optimal presentation of the CARDS and hence maintenance of active signaling (26). In another conceivable scenario, the hydrolysis of ATP might contribute to the tight regulation of the innate immune response by driving the cycling of RLRs between an inactive (RNA-unbound) and an active (RNA-bound) state. In this scenario, posttranslational modifications and/or protein interactions (see section 1.4.1) of the temporarily accessible CARDS finally trap RIG-I and MDA5 in a long-term active state as for steric reasons the modified CRADS cannot be rebound by the ATPase domain. After this step, the activation of downstream signaling pathways via the adaptor protein MAVS can finally occur (60).

### **1.4 RIG-I and MDA5 use the adaptor MAVS to induce cytokine production**

The release or presentation of the CARDS by RNA-bound monomeric RIG-I and filamentous MDA5 is the actual trigger of antiviral signaling pathways. The accessible CARDS are believed to undergo posttranslational ubiquitination catalyzed by particular E3 ligase machineries and/or non-covalent interactions with specific ubiquitin species that qualify the CARDS to enter a homotypic interaction with the signaling adaptor protein MAVS. Subsequent polymerization of the adaptor MAVS transmits the signal via pathways that terminate with the activation of transcription factors operating the production of proinflammatory, antiviral and immunomodulating cytokines. The signaling activity of MDA5 and RIG-I is supposed to be regulated by the third RLR member LGP2 that is not in possession of CARDS and hence unable to signal by itself.

#### **1.4.1 Ubiquitin and the polymerization of MAVS propagate the antiviral signaling cascade**

One of the first events that are believed to go along with the RNA-induced release of the CARDS is their robust ubiquitination mediated by the K63-specific RING-finger E3 ligases TRIM25 (member of the tripartite motif (TRIM) protein family) and Riplet (RING finger protein leading to RIG-I activation). Both enzymes exhibit a PRYSPRY domain that is supposed to bind the first RIG-I CARD once it becomes available upon binding of RIG-I to an RNA-antagonist (see section 1.3.1). The subsequent attachment of K63-linked polyubiquitin chains to specific lysine residues (biochemical studies revealed K 154, 164 and 172 as primary targets) within the second CARD of RIG-I traps the CARDS in their accessible state (63-66) and enhances a homotypic CARD-CARD interaction with the signaling adaptor protein MAVS

(mitochondrial antiviral signaling protein). In an alternative scenario, TRIM25 and Riplet generate a pool of free K63-linked tetraubiquitin chains that interact with the CARDs of RIG-I as soon as they are released. This non-covalent binding of unanchored polyubiquitin chains is thought to promote an oligomerization of the RIG-I receptor molecules and the subsequent CARD-CARD interaction with MAVS (30, 66-68).

MAVS (also known as IPS-1, VISA or Cardif) was identified as an antiviral signaling adaptor protein that acts downstream of RIG-I and MDA5 in 2005. A C-terminal transmembrane domain targets MAVS to the outer mitochondrial membrane, whereas an N-terminal CARD allows interaction with the CARDs of RIG-I and MDA5 (69-73) (see figure 6). In 2011, a striking study provided evidence that RIG-I, in presence of K63-linked tetraubiquitin chains, supplies the initial impetus for a prion-like conversion of monomeric MAVS into functional, fibrillar polymers on the mitochondrial membrane. This conformational switch of MAVS goes along with the activation and propagation of antiviral signaling cascades (see section 1.4.2) (74).

Similar data exist for MDA5. It appears that the CARDs of MDA5 share RIG-I's capability to bind unanchored K63-linked tetraubiquitin chains and that this binding is essential for MDA5 to activate antiviral pathways (67, 68). However, current data suggest that MDA5 CARDs possess an intrinsic oligomerization ability that distinct them from RIG-I CARDs (30). This specific feature and the given proximity of the MDA5 CARDs within RNA-bound filaments are assumed to allow those CARDs to form defined oligomeric patches at the outside of the filaments similar to other death-domain complexes (75). Oligomeric assemblies of the isolated MDA5 CARDs have been shown to be sufficient to induce MAVS polymerization *in vitro*, whereas the monomeric CARDs of RIG-I required the presence of K63-linked tetraubiquitin chains (30). These observations imply differing roles of K63-linked ubiquitin chains in signaling by MDA5 and RIG-I.

Albeit the precise mechanism and the structural details of MAVS binding and activation by RIG-I and MDA5 are still unknown, it is generally accepted that the assembly of a functional MAVS signalosome represents a crucial step for RLRs to propagate antiviral signaling (see figure 6).

## 1.4.2 The activated transcription factors NF- $\kappa$ B, IRF3 and IRF7 drive the cytokine production

Activated MAVS is known to interact with several different downstream signaling proteins. Among these molecules are TRADD (tumor necrosis factor receptor type-1 associated death domain), RIP1 (receptor-interacting serine/threonine-protein kinase 1), FADD (Fas-associated protein with death domain), TRAF6 and TRAF3 (tumor necrosis factor receptor associated factors) (4, 66, 76, 77). All these factors form an

antiviral signaling platform that represents a branching point by mediating downstream signaling to IRF3/7 (interferon regulatory factors) as well as NF- $\kappa$ B (nuclear factor- $\kappa$ B).

TRAF3 mediates the activation of the IKK-related kinases (see NF- $\kappa$ B activation) TBK1 (TANK-binding kinase 1) and IKK $\epsilon$  (inhibitor of  $\kappa$ B kinase  $\epsilon$ ), which are required for the phosphorylation of IRF3 and IRF7. The dimerization of phosphorylated IRFs promote their translocation to the nucleus, where they upregulate the expression of type I interferons (66, 78).

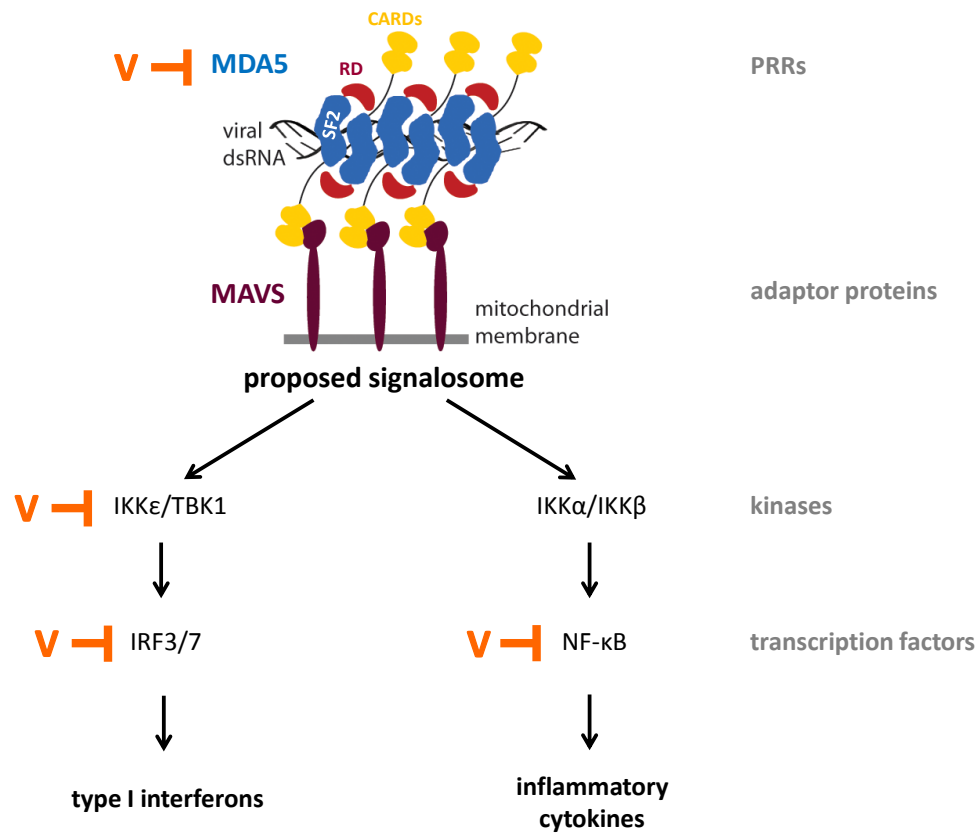
The association of activated MAVS with TRAF6, TRADD, RIP1 and FADD is followed by the caspase-8/10 dependent activation of the I $\kappa$ B kinase (IKK) complex (4, 79). This complex, which comprises the two catalytic subunits IKK $\alpha$  and IKK $\beta$  as well as the regulatory subunit IKK $\gamma$  (also known as NEMO (NF- $\kappa$ B essential modulator)), is required to mediate downstream signaling to NF- $\kappa$ B (80). The five members p65 (RelA), RelB, c-Rel, p50 and p52 of the NF- $\kappa$ B transcription factor family act as homo- or heterodimers. In unstimulated cells, regulatory I $\kappa$ B (inhibitor of  $\kappa$ B) proteins sequester the NF- $\kappa$ B dimers in the cytoplasm by masking their NLS (81, 82). Upon stimulation, the activated IKK complex phosphorylates the I $\kappa$ B proteins at specific serine residues. This phosphorylation primes I $\kappa$ B proteins for K48-linked polyubiquitination that results in degradation by the 26S proteasome. The released NF- $\kappa$ B dimers translocate to the nucleus, where they promote the expression of various proinflammatory cytokines and antimicrobial effector molecules (83, 84).

Type I interferons and other cytokines are rapidly induced during viral infection and are secreted by infected cells. The released cytokines bind distinct surface receptors on target or effector cells. This cytokine signaling causes activation of the transcription factor ISGF3 composed of STAT1, STAT2 (signal transducers and activators of transcription) and IRF9 and allows transcriptional activation of a common set of ISGs (interferon stimulated genes). The ISG encoded proteins mediate antiviral protection by inhibition of the viral lifecycle and stimulatory effects that coordinate the transition to the adaptive immune response during infection. It is this pathway that constitutes the essential role of virus induced cytokines for our survival (85, 86).

### 1.4.3 CARDless LGP2 regulates the antiviral activity of RIG-I and MDA5

LGP2 is in the position to bind dsRNA-antagonists (31, 87), but given the absence of any CARD in LGP2 it cannot transmit this signal. That raises the question which actual function the third RLR member adopts and fueled speculations regarding a regulative role of LGP2 that modulates signaling by RIG-I and MDA5. Indeed, several studies report an inhibitory role of LGP2 in RIG-I signaling. In these studies transient overexpression of LGP2 in living cells inhibited the production of type I interferons and NF- $\kappa$ B-





**Figure 6:** Simplified illustration and schematic overview of MDA5 filament formation, MAVS activation and subsequent antiviral signaling pathways that can be inhibited at multiple levels by the paramyxoviral V protein (highlighted in orange). Adapted from (30, 53, 54, 61, 109,111, 113, 114, 130).

dependent pathways upon Newcastle disease virus or Sendai virus infection (21, 22, 31). This prompted the view that LGP2 acts as a negative feedback regulator of signaling initiated by RIG-I, which is known to sense these paramyxoviruses (see section 1.2.1). In support of this proposed inhibitory role, LGP2 knockout mice displayed enhanced resistance to intranasal infection by vesicular stomatitis virus, a negative strand virus specifically recognized by RIG-I (see section 1.2.1). In addition, the LGP2 deficient mice were observed to exhibit an impaired type I interferon production in response to infection by EMCV, a picornavirus sensed by MDA5 (see section 1.2.1). This finding implies that LGP2 contributes to positive regulation of MDA5 mediated antiviral signaling (88). In contrast, a study using a different strain of LGP2 knockout mice provided evidence that LGP2 facilitates signaling by both MDA5 and RIG-I (19). The function of LGP2 in virus recognition and signaling still remains controversial. It has been suggested that LGP2 might modify virus RNA by removing viral RNA binding proteins, alter the

conformation of RNA, modulate the intracellular localization of or sequester viral RNA to influence the function of RIG-I and MDA5 (18). Many questions remain and further mechanistic studies will be required to resolve the precise action of LGP2 in antiviral signaling.

### **1.5 The paramyxoviral V protein is a potent cytokine antagonist**

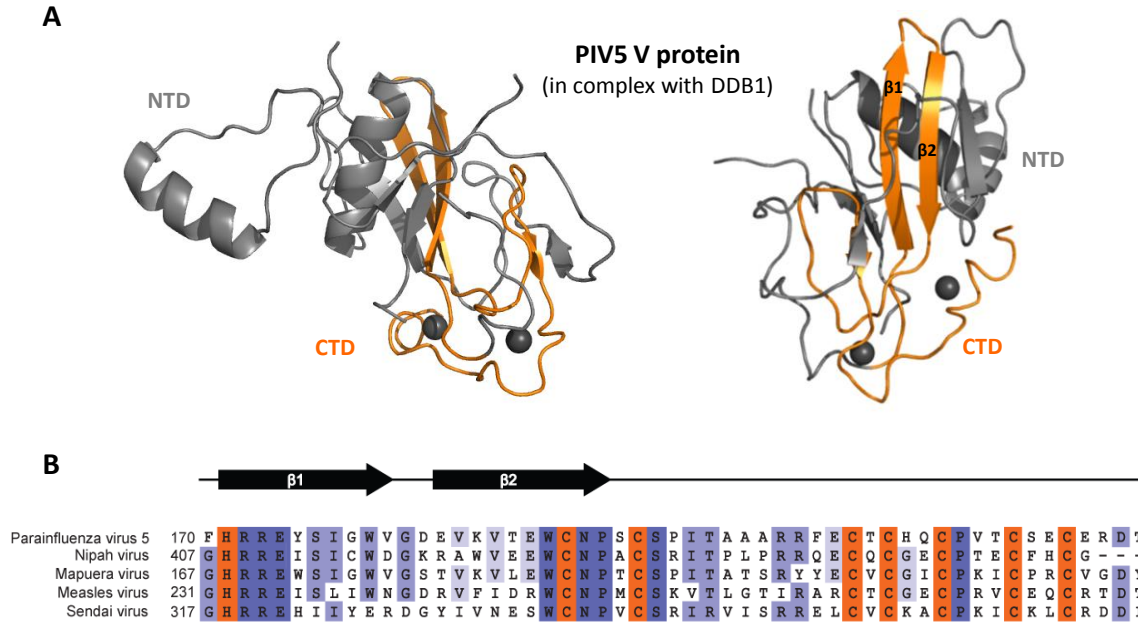
Viruses evolved numerous and various mechanisms to counteract the host immune response (89-97). This section focuses on the V protein that paramyxoviruses produce to antagonize central elements of the innate immune system. This protein severely disturbs the production of as well as signaling by immunomodulating cytokines at several levels, but also attacks the first step of viral ligand sensing by directly targeting MDA5.

#### **1.5.1 The V protein is a non-structural protein required for viral pathogenesis**

The V protein is a non-structural protein that is not incorporated into the virus particle (98), but rather represents a virulence factor (99). The V protein is expressed by the P gene of paramyxoviruses. This gene exhibits two alternative open reading frames that give rise to the structural P and the non-structural C protein. An additional RNA-editing process generates a modified mRNA that is translated into the V protein. As a result of the editing process, the P and the V protein share an identical N-terminal sequence but possess different C-termini (100, 101). The C-terminal domain (CTD) of the V protein is highly conserved among the family of paramyxoviruses and is known to fold into a unique zinc binding motif (see figure 7). In this motif, which shows homology to zinc finger motifs found in many transcription factors and other nucleic acid binding proteins, seven highly conserved cysteines and one histidine coordinate two zinc ions (100, 102-105). The less conserved N-terminal domain (NTD) of the V protein is assumed to be more flexible and partly unfolded. In case of the Parainfluenza virus 5 (PIV5) V protein, the NTD is partially stabilized by the C-terminal zinc binding motif whose two  $\beta$ -strands form a continuous  $\beta$ -sheet with the NTD (see figure 7A). However, the overall structure of the NTD appears to be very transient and is characterized by many flexible loop regions (104, 106). This intrinsic flexibility and structural plasticity might represent the key that allows the V protein to adapt to many different target proteins and hence to function as potent antagonist of the host immune response (see sections 1.5.2 and 4.4).

In 1997, a study using a modified Sendai virus strain unable to produce the V protein showed that this protein is not essential for the replication cycle of the virus *per se*, but maintains viral replication in

later stages of infection. These results suggested a critical role of the V protein in viral pathogenicity (99). Five years later, two studies revealed the V protein as potent interferon antagonist that inhibits both the induction of type I interferons as well as interferon signaling (107, 108). Many mechanistic studies followed and identified several specific targets of the V protein that function at different levels of innate immune pathways (see figure 6).



**Figure 7:** (A) Crystal structure of the Parainfluenza virus 5 (PIV5) V protein in complex with DDB1 (DNA damage-binding protein 1, not shown here). The N-terminal domain (NTD) is shown in grey. The conserved zinc binding motif of the C-terminal domain (CTD) is highlighted in orange and its two bound zinc ions are shown in grey. Adapted from (104). (B) Structure-based alignment of the conserved CTD of selected paramyxoviral V proteins with annotated secondary structure elements. Residue conservation is color-coded according to percentage identity. The highly conserved histidine and cysteine residues, which coordinate the two zinc ions within the CTD, are highlighted in orange.

### 1.5.2 The V protein sabotages the innate immune response at several levels

Inhibition of innate immune signaling cascades by paramyxoviral V proteins at the transcriptional level has been addressed in several studies (see also figure 6). Transcription factors of the NF- $\kappa$ B family usually function as p65/p50 heterodimers. The Measles Virus (MV) V protein has been shown to bind the p65 subunit and to prevent its nuclear translocation. In this way, the V protein is able to suppress p65/p50-mediated NF- $\kappa$ B activity and to impair NF- $\kappa$ B-dependent gene expression (109). Similar results

have been reported for the transcription factor IRF3. The binding of V proteins of different paramyxovirus family members to IRF3 has been revealed to retain the transcription factor in the cytosol (*110, 111*). Moreover, the V protein is assumed to induce degradation of IRF3 (*112*). Another study reported binding and transcriptional inactivation of IRF7 by MV V protein indicating that this transcription factor represents an alternative target of paramyxoviral V proteins (*113*).

At the level of activating kinases, the paramyxoviral V protein has been revealed to inhibit IRF3/IRF7 activation in an indirect manner by acting as a decoy substrate of IKK $\epsilon$ /TBK1 (*114*). V proteins are directly phosphorylated by this IRF3/7 kinase. Thus, V proteins are proposed to interfere with IRF3/7 phosphorylation by mimicking and competing out IRF3/7 as IKK $\epsilon$ /TBK1 substrates.

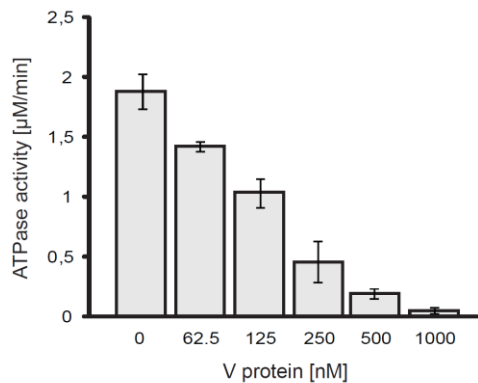
NF- $\kappa$ B, IRF3/7 and IKK $\epsilon$ /TBK1 represent targets that are essential for cytokine induction (see section 1.4.2). But the inhibition by paramyxoviral V proteins goes beyond this and also affects signal transduction by cytokines. The transcription regulators STAT1 and STAT2 are key components of the paracrine and autocrine cytokine signaling system (see section 1.4.2) and are subjected to direct and indirect inhibition by V proteins. The V proteins of Nipah virus, Hendra virus as well as MV are known to bind and sequester STAT1 and STAT2 in the cytoplasm to prevent their nuclear accumulation. The MV V protein has been shown to bind in addition the third member IRF9 of the ISG3 complex. The Nipah virus V protein inhibits cellular response to cytokines also by preventing activating tyrosine phosphorylation of STAT proteins, while the MV V protein is not able to interfere with this step (*115-117*). In contrast to these direct inhibition strategies, many paramyxoviruses have evolved indirect mechanisms to disturb signaling by STAT proteins including the proteasomal degradation of these proteins. Mumps virus, PIV5 and human PIV2 use their V protein to recruit STAT1, STAT2, Cul4A (cullin RING ubiquitin ligase 4A) and its substrate recruiting adaptor protein DDB1 (DNA-damage-binding protein-1) to assemble STAT-degradation complexes (also called VDCs = V protein dependent degradation complexes). The DDB1-Cul4A ubiquitin ligase machinery is usually involved in the regulation of DNA repair, DNA replication and transcription. Binding of the V protein to DDB1 manipulates the substrate specificity of the ubiquitin ligase complex and trains it for STAT-degradation (*104, 118-126*) (see also section 4.4).

The numerous studies investigating STAT inhibition by paramyxoviruses demonstrate its functional diversity and have taught us that the mechanistic details of V protein-dependent evasion strategies can highly depend on the virus species and its host (*95, 127-129*).

### 1.5.3 The V protein specifically binds and inhibits MDA5

Beside MDA5, many other immune receptors (e.g. RIG-I, TLR 3, 4, 7 and 8) use IKK $\alpha$ /IKK $\beta$  or IKK $\epsilon$ /TBK1 as terminal kinases to activate NF- $\kappa$ B or IRF transcription factors. Paramyxoviral V proteins have the potential to inhibit all these signaling pathways by targeting most of these shared key components. In addition, paramyxoviruses are in the position to specifically and severely affect the MDA5 signaling cascade at the very first step by targeting MDA5 itself. The direct inhibition of MDA5 by paramyxoviral V proteins was firstly described in 2004 (130). This initial study revealed that V proteins of different paramyxoviruses prevent the activation of both NF- $\kappa$ B and IRF3/7 by binding to MDA5. Further characterization of this interaction showed that the conserved zinc binding motif of the CTD shared by paramyxoviral V proteins (see section 1.5.1) is sufficient to mediate MDA5 binding and requires structural integrity of this motif (110, 131-133). A subsequent study investigating a wide variety of 13 paramyxoviruses (namely PIV5, human PIV2, porcine rotavirus (PoRV), Newcastle disease virus (NDV), mumps, measles, Menangle, Hendra, Nipah, Mapuera, Sendai, Salem and Tioman virus) showed that all tested V proteins target MDA5 via their conserved cysteine-rich CTD in a direct mode (134). Despite the high sequence similarity between RLR family members, the tested V proteins were found to fail at binding of RIG-I, but were able to interact with LGP2 (134-136). Detailed mapping of the interaction sites identified the C-terminal RecA-like fold within the ATPase domain of MDA5 and LGP2 as common target region for paramyxoviral V proteins (135, 136). In addition, some paramyxovirus V proteins are able to target further regions of MDA5, which include for instance the N-terminal RecA-like fold (136). In contrast to these virus specific interactions, the binding involving the conserved CTD of V proteins and the C-terminal RecA-like fold of MDA5 (and LGP2) appears to be universal. The described data point toward a conserved interaction interface that allows paramyxoviruses to specifically inhibit MDA5, but not RIG-I signaling initiation.

*In vitro* experiments that investigated the impact of V protein binding on MDA5 function demonstrated that V protein interaction affects ATP hydrolysis mediated by MDA5. Titration of purified V protein into a solution containing purified and functional MDA5 was shown to cause a complete shutdown of the MDA5 ATP hydrolysis activity (see figure 8) (57, 135, 137). Reports regarding the interference of V protein interaction with MDA5 RNA binding activity are contradicting. A study proposing that V proteins block the RNA binding capability of the MDA5 ATPase domain (136) is in opposition to data showing that V protein interaction does not interfere with RNA binding by MDA5 (137). However, no study provided evidence that paramyxoviral V proteins are able to disturb the cooperative filament assembly of MDA5 along dsRNA.



**Figure 8:** *In vitro* experiment demonstrating the dose-dependent shutdown of MDA5 ATP hydrolysis activity by the paramyxoviral V protein (57, 137).

## 1.6 Objectives

The aim of this work was the crystallization of a complex between MDA5 and the paramyxoviral V protein to derive a first molecular mechanism of viral interference with RLRs. Structural determination of this complex was assumed to provide important mechanistic details regarding recognition and inhibition of MDA5 by paramyxoviral V proteins. The identification of the C-terminal RecA-like fold as common target site for V proteins (135,136) (see section 1.5.3) implies the possibility that V protein interaction might interfere with the formation of a mature dsRNA- and active ATP-binding site upon ligand recognition by MDA5 (see sections 1.3.1 and 1.3.2). Structural investigation of the viral inhibitor protein bound to its target MDA5 was expected to reveal how exactly the V protein affects the ATPase activity of MDA5 and to answer the question whether and in what manner the V protein interferes with the cooperative RNA binding mode of MDA5. These questions address major aspects as both, ATPase activity as well as proper RNA binding, have been shown to be crucial for antiviral signal propagation by MDA5 (30) (see section 1.3.3). Structural knowledge at the atomic level and the deduction of a mechanistic inhibition model were assumed to pave the way for precise and verifying biochemical studies examining the impact of V protein binding on the function of MDA as viral RNA receptor. Moreover, the detailed determination of the interaction interface between MDA5 and the V protein was assumed to explain, why the inhibitor protein specifically recognizes and binds the two RLR family members MDA5 and LGP2, whereas the highly related RIG-I molecule is able to evade V protein binding. To this end, structural analysis of the MDA5:V protein complex was expected to contribute to our general understanding of RLR function and to give significant insights into virus-host interactions and viral immune evasion strategies.

## 2. Materials and Methods

### 2.1 Materials

All chemicals used in this work were of the highest available grade obtained from Roth, Merck or Sigma-Aldrich, unless otherwise stated. Oligonucleotide primers for polymerase chain reaction (PCR) were obtained in “high purity salt free” (HPSF<sup>®</sup>) form from Metabion or Biomers. Enzymes for molecular biology were obtained from New England Biolabs and Stratagene. Crystallization screens and tools were from Hampton Research, NeXtal (QIAGEN) and Jena Bioscience. Electrophoretic mobility shift assays were performed with HPLC-grade RNA oligonucleotides from biomers. Synthetic poly(I:C) RNA was purchased from Sigma-Aldrich for activity assays and from The Midland Certified Reagent Company for negative-stain electron microscopy.

### 2.2 Methods

The sections 2.2.1 and 2.2.2 list basic methods employed throughout the work. Further methods are stated in chronological order according to their declaration in chapter 3.

#### 2.2.1 General molecular biology methods

##### 2.2.1.1 Cloning

DNA fragments encoding the desired protein constructs were amplified from plasmids or cDNA containing the sequence of interest using the Phusion High-Fidelity Master Mix (New England Biolabs) and standard oligonucleotide primers (Metabion). Amplified DNA fragments were purified from preparative agarose gels using the NucleoSpin Gel and PCR Clean-up kit from Macherey-Nagel and subjected to restriction enzyme (New England Biolabs) digest. Ligation reactions of digested and purified insert DNA with vector DNA were performed with T4 DNA Ligase (New England Biolabs) and incubated for 1 h at room temperature. Subsequent transformation of *Escherichia coli* XL1-Blue (K12) was performed as described in section 2.2.1.3. All cloned constructs were verified by DNA sequencing.

##### 2.2.1.2 Site-directed mutagenesis

PCR for site-directed mutagenesis was performed using the Pfu Ultra DNA Polymerase (Stratagene) and standard oligonucleotide primers (Biomers). (Dam)-methylated, non-mutated template DNA was digested

by 20 U DpnI (New England Biolabs) for 2 h at 37°C. Subsequent transformation of *Escherichia coli* XL1-Blue (K12) was performed as described in section 2.2.1.3. All mutants described in this work were generated by this standard protocol and constructs were verified by DNA sequencing.

### 2.2.1.3 Heat shock transformation of *Escherichia coli* XL1-Blue (K12)

Chemically competent *Escherichia coli* XL1-Blue (K12) cells were mixed with plasmid DNA and incubated for 20 min on ice. After heat-shocking for 2 min at 42°C, the cells were incubated for 1 min on ice. The cells were supplemented with 1 ml LB-medium and allowed to grow for 45 min at 37°C before spreading on LB-plates containing the appropriate antibiotic for selection.

### 2.2.1.4 Isolation of plasmid DNA from *Escherichia coli* XL1-Blue (K12)

Single colonies were picked and inoculated into 4 ml LB-media supplemented with the appropriate antibiotic (see table 1). The cultures were allowed to grow overnight at 37°C and 190 rpm and harvested the next day by centrifugation (5000 g for 5 min). Plasmid DNA was isolated using the NucleoSpin Plasmid Quick Pure kit from Macherey-Nagel according to the instructions of the manufacturer.

### 2.2.1.5 Determination of nucleic acid and protein concentrations

Nucleic acid and protein concentrations were quantified using the Thermo Scientific NanoDrop™ 1000 Spectrophotometer (PeqLab). The measured protein absorbance at 280 nm was adjusted using the molar extinction coefficient for the corresponding amino acid sequence calculated by ExPASy.

## 2.2.2 General bioinformatics methods

### 2.2.2.1 Protein sequence alignments and secondary structure predictions

Multiple protein sequence alignments were built with Jalview/Muscle. Protein secondary structures were predicted based on sequence alignments with Jalview/JNet.

### 2.2.2.2 Structure visualization and analysis

Visualization and analysis (e.g. three-dimensional superposition, distance measurements, analysis of flexible and unstructured protein regions or crystal contacts) of structures were achieved with PyMol.



### 2.2.3 Small scale protein testexpression

*Escherichia coli* Rosetta (DE3) cells (Novagen) were transformed with the desired plasmid DNA according to the protocol described in 2.2.1.3 and allowed to grow overnight at 37°C and 190 rpm in LB-media supplemented with antibiotic(s) for selection (see table 1). The overnight pre-culture was inoculated in a ratio of 1:100 into 60 ml LB-medium containing the appropriate antibiotic(s). Small scale testexpression cultures were incubated overnight at defined test conditions (induction of gene expression with either 0.5 mM or 0.1 mM IPTG (final concentration) at OD<sub>600</sub> and 20°C) and harvested by centrifugation (5000 g for 5 min). The bacterial pellets were suspended in 5 ml suitable lysisbuffer (see section 2.2.5 and table 3) that contained 1 mg/ml lysozyme. After 20 min incubation on ice, the samples were additionally sonicated at 4°C using a Branson Sonifier 250 (2x 30 s at an intensity of 80% and at an output power level of 2) to ensure complete bacterial cell lysis. To permit identification of non-soluble protein fractions, a 5 µl aliquot of the disrupted cells was taken and mixed with SDS containing protein loading buffer for later analysis. The remaining sample was centrifuged for 30 min at 48 000 g and 4°C. To permit later analysis of soluble protein fractions, a 5 µl aliquot of the supernatant was taken and mixed with SDS containing protein loading buffer. Glutathione-coated cellulose beads (Roth) or Ni-NTA agarose beads (Macherey-Nagel) were equilibrated with washingbuffer (lysisbuffer without lysozyme, see section 2.2.5 and table 3) and a 1:1 bead suspension was prepared. 50 µl of the bead suspension was added to 1 ml of the supernatant and the samples were incubated for 1 h at 4°C on a rotating wheel. The beads were washed four times with 1 ml washingbuffer. To elute protein from beads, the bead fraction was mixed with SDS-containing protein loading buffer and incubated together with the previously taken samples of disrupted cells and supernatant for 10 min at 90°C. To separate beads from eluted proteins, the samples were centrifuged for 1 min at 16 000 g. The samples (1 µl of disrupted cells and supernatant respectively, 5 µl of bead eluate) were loaded on a 15% SDS-polyacrylamide gel and proteins were separated by electrophoresis according to Laemmli. Coomassie staining of the gel allowed analysis of protein yield and solubility.

**Table 1:** Antibiotic stock solutions (used in 1:1000 dilution).

Antibiotic	Concentration	Solvent
Ampicilin (Na-salt)	100 mg/ml	water
Kanamycin	50 mg/ml	water
Chloramphenicol	50 mg/ml	ethanol

---

#### 2.2.4 Large scale protein expression

Sequences encoding full length mmMDA5 (1-1025aa) and ggMDA5 (4-1001aa) were cloned into a modified pET-28a(+) vector (Novagen) carrying a TEV cleavage site. Sequences encoding the SF2 ATPase domain of hsMDA5 (304-890aa), mmMDA5 (305-890aa), ssMDA5 (303-887aa), btMDA5 (303-885aa), ggMDA5 (297-864aa) and tgMDA5 (363-935aa), the RDless version of mmMDA5 (1-890) and the CARDless version of hsMDA5 (304-1017) were cloned into the pET-21a(+) vector (Novagen). The oligonucleotides used for cloning of these MDA5 constructs are listed in table 2. A modified pGEX vector containing the sequence of full-length parainfluenza virus 5 (PIV5) V protein was received from Ning Zheng (University of Washington). According to the testexpression results, expression of all constructs was carried out in *Escherichia coli* Rosetta (DE3) (Novagen) overnight at 20°C by induction with 0.1 mM IPTG. The cultures were harvested by centrifugation (5000 g for 15 min) and the bacterial pellets were frozen at -80°C.

**Table 2:** Oligonucleotide list – MDA5 constructs.

Primer name	5'-3' sequence
pET28 mmMDA5 M1 NheI forward	catggctagcatgtcgattgtctgtctgca
pET28 mmMDA5 D1025 BamHI reverse	gcgggatccttaattctcatcactatacaagcagt
pET28 ggMDA5 E4 NheI forward	catggctagcgcagtgccgagacgagcgc
pET28 ggMDA5 D1001HindIII reverse	gtacaagcttttaattctcatcactgaaggacaat
pET21 hsMDA5 P304 NheI forward	catggctagccagaactccagctcaggc
pET21 hsMDA5 R890 XhoI reverse	catgctcgcagtcctcttggttttcatttttttc
pET21 mmMDA5 P305 NheI forward	catggctagccagaactgcagctcaggc
pET21 mmMDA5 R890 NotI reverse	aaggaaaaaagcgccgcctcttttgactttcatttttttcc
pET21 ssMDA5 P303 NheI forward	catggctagccagagctgcattcaggc
pET21 ssMDA5 R887 XhoI reverse	catgctcgcagtcctcttgattttcatttttttcca
pET21 btMDA5 P303 NheI forward	catggctagccagagctgaattcaggcc
pET21 btMDA5 R885 XhoI reverse	catgctcgcagtcctcttggttttcatttttttcc
pET21 ggMDA5 P297 NheI forward	catggctagccagatctcaccctgagaga
pET21 ggMDA5 R864 SalI reverse	cagtgtcgactctctttgccttcattgtttttc
pET21 tgMDA5 P363 NheI forward	catggctagccagatctgaccctgagagat
pET21 tgMDA5 R935 SalI reverse	cagtgtcgactctcatcacttcattgtttttcc
pET21 mmMDA5 M1 NheI forward	catggctagcatgtcgattgtctgtctgca
pET21 mmMDA5 R890 NotI reverse	aaggaaaaaagcgccgcctcttttgactttcatttttttcc
pET21 hsMDA5 P304 NheI forward	catggctagccagaactccagctcaggc
pET21 hsMDA5 E1017 XhoI reverse	catgctcgcagttctgaatagcaagattgggaa

### 2.2.5 Large scale purification of MDA5 and PIV5 V proteins and their complexes

Full length and RDless MDA5 proteins were purified by nickel affinity chromatography (Ni-NTA agarose beads, Macherey-Nagel), anion-exchange chromatography (HiTrap Q HP column, GE Healthcare) and final size exclusion chromatography (Superdex200 26/60 or 16/60 column, GE Healthcare). A similar protocol was used for the purification of MDA5 ATPase domains and CARDless hsMDA5, but cation-exchange (HiTrap S HP column, GE Healthcare) instead of anion-exchange chromatography was performed. Ahead of size exclusion chromatography, CARDless hsMDA5 was additionally cleaned via a heparin column (HiTrap Heparin HP, GE Healthcare). The PIV5 V protein was batch-purified using glutathione-coated cellulose beads (Roth). The GST-tag of the construct was removed by on-beads-cleavage using His<sub>6</sub>-tagged TEV-protease. The TEV protease was rebound to Ni-NTA agarose beads and further purification of the PIV5 V protein was achieved by size exclusion chromatography (GE Healthcare, Superdex75 26/60). For complex formation, proteins were mixed in a 1:1.5 molar ratio (MDA5/PIV5 V protein) in a final complex buffer (see table 3) and incubated for 1 h on ice. A stoichiometric 1:1 complex was subsequently purified by size exclusion chromatography (Superdex200 26/60 or 16/60 column, GE Healthcare). Lysis of bacterial cells was achieved with a Branson Sonifier 250 (2x 1 min at an intensity of 40% and at an output power level of 8) at 4°C. The purification steps were performed on the Äkta FPLC, Äkta basic or Äkta purifier system (GE Healthcare). Table 3 lists and states all buffers used for the different purification steps. Purification of the protein was monitored at all steps by SDS-PAGE according to Laemmli and Coomassie staining. Concentration of protein solutions was achieved with Amicon Ultra Centrifugal Cellulose Filters (Millipore).

**Table 3:** Protein purification buffers.

Buffer	Composition	final pH
Lysisbuffer	500 mM NaCl, 50 mM Tris, 20 mM imidazole, 2 mM MgCl <sub>2</sub> ,	7.8
Ni-NTA affinity chromatography	10 $\mu$ M ZnCl <sub>2</sub> , 10 mM $\beta$ -mercaptoethanol	
Lysisbuffer	500 mM NaCl, 50 mM Tris, 10 $\mu$ M ZnCl <sub>2</sub> ,	7.8
Glutathione affinity chromatography	10 mM $\beta$ -mercaptoethanol	
Washingbuffer	100 mM NaCl, 50 mM Tris or Bis-Tris, 20 mM imidazole,	8.5 or 6.5
Ni-NTA affinity chromatography	2 mM MgCl <sub>2</sub> , 10 $\mu$ M ZnCl <sub>2</sub> , 10 mM $\beta$ -mercaptoethanol	
Washingbuffer	150 mM NaCl, 50 mM Tris, 10 $\mu$ M ZnCl <sub>2</sub> ,	7.8
Glutathione affinity chromatography	10 mM $\beta$ -mercaptoethanol	
Elutionbuffer	100 mM NaCl, 50 mM Tris or Bis-Tris, 200 mM imidazole,	8.5 or 6.5
Ni-NTA affinity chromatography	2 mM MgCl <sub>2</sub> , 10 $\mu$ M ZnCl <sub>2</sub> , 10 mM $\beta$ -mercaptoethanol	
Low salt buffer	100 mM NaCl, 50 mM Tris, 2 mM MgCl <sub>2</sub> , 10 $\mu$ M ZnCl <sub>2</sub> ,	8.5
Anion-exchange chromatography	10 mM $\beta$ -mercaptoethanol	
High salt buffer	1 M NaCl, 50 mM Tris, 2 mM MgCl <sub>2</sub> , 10 $\mu$ M ZnCl <sub>2</sub> ,	8.5
Anion-exchange chromatography	10 mM $\beta$ -mercaptoethanol	
Low salt buffer	100 mM NaCl, 50 mM Bis-Tris, 2 mM MgCl <sub>2</sub> , 10 $\mu$ M ZnCl <sub>2</sub> ,	6.5
Cation-exchange chromatography	10 mM $\beta$ -mercaptoethanol	
High salt buffer	1 M NaCl, 50 mM Bis-Tris, 2 mM MgCl <sub>2</sub> , 10 $\mu$ M ZnCl <sub>2</sub> ,	6.5
Cation-exchange chromatography	10 mM $\beta$ -mercaptoethanol	
Size exclusion chromatography buffer	150 mM NaCl, 20 mM Tris, 10 $\mu$ M ZnCl <sub>2</sub> ,	7.8
PIV5 V protein	5 mM DTT	
Size exclusion chromatography buffer	150 mM NaCl, 20 mM Tris, 5 mM MgCl <sub>2</sub> , 10 $\mu$ M ZnCl <sub>2</sub> ,	7.8
MDA5 and complex formation	5 mM DTT	

### 2.2.6 ATPase assays

ATPase assays were performed in a 10  $\mu$ l reaction volume containing 100 mM Tris pH 7.9, 150 mM NaCl, 5 mM MgCl<sub>2</sub>, 5  $\mu$ M ZnCl<sub>2</sub>, 1 mM DTT, 0.1 mM ATP, 20 nM  $\gamma$ -<sup>32</sup>P-ATP (3000 Ci/mmol, Hartmann Analytic) and 50 nM poly(I:C) RNA (Sigma-Aldrich). After equilibration to 37°C, the reaction was started by addition of purified mmMDA5 (final concentrations 1  $\mu$ M, 2.5  $\mu$ M, 5  $\mu$ M, 10  $\mu$ M and 15  $\mu$ M). After 20 min incubation at 37°C, 1  $\mu$ l of the reaction mix was spotted onto a thin layer chromatography polyethyleneimine cellulose plate (Merck). The TLC plate was developed in 0.5 M LiCl / 1.0 M formic acid and subsequently air-dried. The amount of educt (ATP) and product (liberated  $\gamma$ -<sup>32</sup>P) was visualized with a phosphorimager (GE Healthcare).

### 2.2.7 Static light scattering (SLS) measurements

Static light scattering represents a method to determine the absolute molar mass of a protein in solution and permits therefore the detection of multimeric states in solution. Static light scattering is different from dynamic light scattering which allows the calculation of the hydrodynamic radius of a molecule from its diffusion coefficient. The principle of static light scattering is based on the induction of an oscillating dipole within a macromolecule when monochromatic laser light impinges on it, resulting in re-radiation of light. The intensity of this scattered light depends among other factors (e.g. polarizability of the molecule/strength of the induced dipole) on the concentration of the macromolecule in solution in a proportional manner. The intensity measurement of the scattered light allows not only the concentration determination of a macromolecule in solution. It also provides information whether a macromolecule adopts a multimeric or monomeric state in solution given the direct proportionality of the light intensity to the molar mass of the scattering molecule. Scattered light originating from protein monomers is characterized by incoherence, because the random Brownian motion of the monomers separated by solvent molecules induces randomness in the phase of the radiated light. In case of dimer formation two monomers of the protein move together in solution which results in a defined phase relationship of the light scattered by the two linked monomers and for this reason in coherence. This effect doubles the intensity of the radiated light, although the concentration of the protein solution is not changed. In this way static light scattering represents a technique to determine the presence of higher multimeric states of a given protein or the stoichiometry of a protein complex in solution.

Static light scattering measurements were performed by Claire Basquin (biophysics facility, Max Planck Institute of Biochemistry, Martinsried). A Superdex200 10/300 column (GE Healthcare) was connected in-line with the following detectors:

1. UV-detector (monitor UV 900 as part of the ÄKTA purifier, wavelengths: 230 nm, 260 nm and 280 nm).
2. Light scattering detector (miniDAWN Treos, Wyatt Technology; allows multiangle static laser light scattering; possesses three detectors (at 45°, 90° and 135°) which measures the relative intensity of scattered light as a function of the angle between the detector and the incident laser beam direction).
3. Refractive index detector (OPTILAB rEX, Wyatt Technology; determines sample mass concentration by measuring the refractive index difference between the sample solution and the reference fluid by utilizing an array of 512 photodetectors).

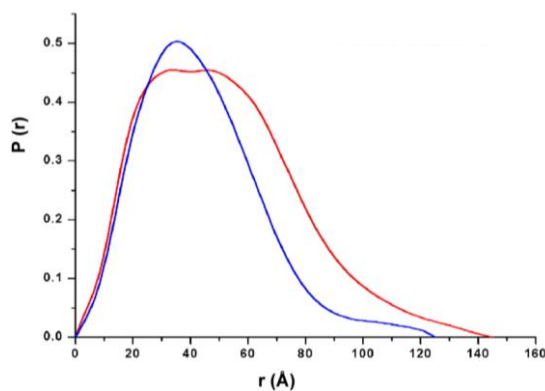
The system was equilibrated with complex formation buffer (see table 3). Approximately 300  $\mu$ M purified protein was applied on the Superdex200 10/300 column and static light scattering measurements

were performed at a flow rate of 0.4 ml/min. UV absorption was monitored by the Unicorn software and the molecular weight was determined by the AstraV software provided by Wyatt Technology.

### 2.2.8 Small-angle X-ray scattering (SAXS) of protein solutions

This X-ray scattering technique measures the distribution of intramolecular distances in macromolecules in solution and is therefore able to provide information about the shape and size of non-crystalline proteins in a buffer of choice. SAXS has no limitation regarding the protein size that can be studied, but the structures obtained by *ab initio* modeling are of relatively low resolution (maximum 15Å). However, the pair distance distribution function, which describes the paired set of all distances between all electrons within a protein, represents a powerful tool to detect conformational changes within a protein in solution. The blue curve in figure 9 shows a symmetric and parabolically shaped distance distribution function characteristic of a globular and compact protein with one single major center of mass. This shape is lost in case of a protein whose conformation is defined by more than one large center of mass (red curve in figure 9).

SAXS data were collected at EMBL beamline X33 (DESY, Hamburg, Germany). Protein samples after size exclusion chromatography in three different concentrations (2 mg/ml, 5 mg/ml and 10 mg/ml) and the running buffer were used for data collection of protein and buffer, respectively. The buffer-corrected scattering curves showed no signs of aggregation and were processed using PRIMUS from the ATSAS package (138) as described. A set of 16 independent *ab initio* models was calculated using GASBOR/DAMMIF (139, 140). Conversion of the final beadmodels obtained by DAMAVER averaging (141) into electron density shapes was done using SITUS (142). All figures were prepared with UCSF Chimera.



**Figure 9:** Pair distance distribution functions of a globular and compact protein with one major centre of mass (blue curve) and of a protein with at least two large centres of mass (red curve) (26).

### 2.2.9 Reductive methylation

MDA5 purification was performed according to the protocols described in 2.2.5, but the final size exclusion chromatography buffer contained 20 mM HEPES instead of Tris. Purified MDA5 proteins were reductively methylated according to a standard protocol (143). After complete methylation, a second gel filtration step was performed to remove reaction contaminants and a third gel filtration step to form a complex between reductively methylated MDA5 and purified PIV5 V protein (see section 2.2.5).

### 2.2.10 Large scale purification of Fab fragments and triple complex formation

Monoclonal antibodies recognizing MDA5 or the PIV5 V protein were produced in collaboration with Dr. Elisabeth Kremmer (Institute for Molecular Immunology, Helmholtz Centre Munich). Purified PIV5 V protein, full length mmMDA5 or ggMDA5 were used as antigens for the immunization of mice. Hybridoma cell lines, which produce IgG antibodies able to recognize epitopes of the antigen, were identified and preselected by the laboratory of Dr. Elisabeth Kremmer. The supernatants of the positively tested hybridoma cell lines were used for pulldown assays (performed with protein G beads) to select high affinity binders that do not interfere with the complex formation between MDA5 and PIV5 V protein. Large scale production and HPLC purification of the selected antibodies expressed by the hybridoma clones GGS/M 4A9, GGS/M 12G5, GGSV 12C10 and GGSV 16B5 was performed by the laboratory of Dr. Elisabeth Kremmer. The purified IgG antibodies were provided in 1x PBS pH 7.0 and fragmentation was performed at a concentration of 20 mg/ml using immobilized papain cross-linked to agarose beads (Thermo Scientific) according to the instructions of the manufacturer. Fab and Fc fragments were separated by size exclusion chromatography (Superdex200 16/60 column, GE Healthcare) using 1x PBS pH 7.5 as running buffer. The isolated Fab fragments were concentrated and added in molar excess to purified MDA5:PIV5 V protein complexes. After 1 h incubation on ice, a stable triple complex containing MDA5, PIV5 V and Fab protein in a 1:1:1 stoichiometry was separated from free Fab fragments by size exclusion chromatography using complex formation buffer (see table 3) supplemented with 0.5 mM DTT (to ensure stability of MDA5 proteins) as running buffer.

### 2.2.11 Crystallization of the ssMDA5:PIV5 V protein core complex

For crystallization, residues 641-665 of the ssMDA5 SF2 ATPase domain construct and residues 55-79 of the PIV5 V protein construct (see section 2.2.4) were replaced with SGSGS and GSGSGSGSGS, respectively. Complex formation of both proteins was performed as described and freshly purified complex was directly used for crystallization. Crystals were obtained by an in-drop proteolysis approach.

To purified complex (10 mg/ml) 1:1000 (w/w) 1mg/ml trypsin (purchased from calbiochem and dissolved in 1mM HCl and 2 mM CaCl<sub>2</sub>) was added and the mixture was crystallized by sitting-drop vapour diffusion in crystallization solution (100 mM Hepes pH 7.5, 20% (w/v) PEG 1500). Crystals appeared within 1 day at 20°C. Crystals of selenomethionine-substituted ssMDA5 protein (expressed in Rosetta B834 (DE3) (Novagen) and SelenoMet<sup>TM</sup> Medium plus Nutrient Mix (Molecular Dimensions)) in complex with unlabeled PIV5 V protein were obtained by streak seeding using native crystal material and hanging-drop vapour diffusion (100 mM Hepes pH 7.5, 18% (w/v) PEG 1500). Crystals were cryoprotected by soaking of the crystals in 20 % 2,3-butanediol in crystallization solution and flash-frozen in liquid nitrogen.

### 2.2.12 Structure determination

#### 2.2.12.1 Principles of X-ray structure analysis

X-ray crystallography allows the structure determination of macromolecules such as proteins at the atomic level. The crystal (in this case a protein crystal) rotating in an incident X-ray beam functions as a diffractor and the diffracted X-rays are recorded as reflections. Each diffracted X-ray results in one reflection and all recorded reflections together create a unique diffraction pattern in a three-dimensional space called the “reciprocal space”. In this reciprocal space the location of each reflection is defined by three indices  $h, k$  and  $l$  representing three-dimensional coordinates.

The reciprocal space stands vis-à-vis a second coordinate system that is represented by the unit cell of the crystal and termed real space. A crystalline lattice is built up by the repeating arrangement of equal unit cells in which an atom's position is given by the fractional coordinates  $x, y$  and  $z$ . Each single reflection in the reciprocal space can be ascribed to the diffractive contributions from all atoms in the unit cell. Each diffracted X-ray recorded as a reflection  $hkl$  can therefore be described as the sum of multiple wave equations (one equation for each diffractive contributor in the unit cell). The computed sum for a given reflection  $hkl$  is called the structure factor  $\mathbf{F}_{hkl}$ .

The actual diffractors in the unit cells are not the atoms but their clouds of electrons. An X-ray impinging on an electron forces the electron to oscillate in the same frequency as the incoming X-ray wave. This causes an emission of scattered X-rays of identical wavelength in all directions of space. The amplitude of the scattered waves is determined by two parameters, the number of electrons and the scattering angle. This type of scattering is known as normal scattering and also called elastic scattering, Thomson scattering or scattering without energy loss.



The recorded diffraction pattern should therefore contain the information of the electron density distribution of the molecules in the crystal. The electron density  $\rho$  at every position  $x,y,z$  in the unit cell is given by following equation:

$$\rho(x, y, z) = \frac{1}{V} \sum_{hkl} |F_{hkl}| e^{i\alpha(hkl)} e^{-2\pi i(hx+ky+lz)} \quad (1)$$

$\rho(x,y,z)$  = electron density at position  $x,y,z$

$V$  = volume of the unit cell

$|F_{hkl}|$  = structure factor amplitude of the reflection  $hkl$

$\alpha$  = phase angle  $\alpha$  of the reflection  $hkl$

$e$  = Euler's number

$i$  = imaginary unit

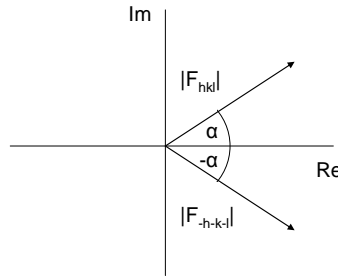
This function connects the reciprocal space of the recorded diffraction pattern with the electron density of the molecules in the real space via the structure factor  $F_{hkl}$ . The graph of this function is the contour map of the electron density of the molecules within the unit cell. The goal is to obtain all unknowns of this equation to compute the electron density map of the desired molecule.

Representing a wave equation, the structure factor  $F_{hkl}$  has to specify three parameters: frequency, amplitude and phase. The recorded diffraction pattern exhibits solely the intensity  $I_{hkl}$  for each reflection  $hkl$ , but can be used for the calculation of the structure factor amplitude  $|F_{hkl}|$  which is proportional to the square root of the measured intensity  $I_{hkl}$ . The frequency of the emitted X-rays is identical with the frequency of the impinging X-ray beam in case of normal scattering. The inverse connection between the reflection distances on the detector and the unit cell dimensions allows the determination of the unit cell volume  $V$  in equation (1) from reflection spacings. The only additional information that is required is the phase angle  $\alpha$ , whereby each reflection  $hkl$  has its own specific phase. The information of the phase angle is lost during the diffraction experiment and has to be obtained in another way. Several possibilities exist to solve the phase problem (= phasing):

- direct solution of the Patterson map (in case of small molecules up to approximately 1000 atoms and high resolution)
- isomorphous replacement (single isomorphic replacement (SIR) or multiple isomorphic replacement (MIR) by e.g. soaking heavy metal ions into a crystal (= generating a heavy metal derivative))
- anomalous dispersion (single wavelength anomalous dispersion (SAD) or multiple wavelength anomalous dispersion (MAD) by using e.g. Se-methionine substituted protein crystals)
- any combination of the previous methods such as SIRAS or MIRAS (SIR and MIR with anomalous scattering)
- molecular replacement (MR) by utilizing a homologous searching model

2.2.12.2 Principles of SAD phasing

Normal scattering or elastic scattering results in the so called Friedel's law stating that members of a Friedel pair of reflections have equal amplitude and opposite phase (see figure 10). Friedel pairs are related by the inversion through the origin and therefore exhibit negated indices ( $hkl$  and  $-h-k-l$ ). When Friedel's law is valid, the atomic scattering factor or atomic form factor  $f_0(\theta)$  is a scalar function and depends only on the scattering angle  $\theta$  and can be treated as a real number.



**Figure 10:** The members of a Friedel pair exhibit equal amplitude and complementary phase when Friedel's law is valid:  $|F_{hkl}| = |F_{-h-k-l}|$ ;  $\alpha_{hkl} = -\alpha_{-h-k-l}$  (adapted from (144)).

Some atoms can function as anomalous scatterers since the energy of the X-ray photons is close to a transition state of their inner shell electrons. In this case the incident X-radiation is elastically scattered but in addition partially absorbed to promote an electron from an inner shell to a higher shell. The excited electron is not able to sustain the higher energy level and falls back to a lower shell resulting in the emission of X-radiation (known as fluorescence). As consequence of the absorption of X-radiation, the atomic form factor is not longer a scalar function but becomes a vectorial quantity consisting of three components: the normal scattering term  $f_0(\theta)$  which is dependent on the scattering angle, the dispersive term  $f'(\lambda)$  and the absorption term  $f''(\lambda)$  that are not dependent on the scattering angle but on the wavelength. The full atomic form factor can be described as a complex number:

$$f(\theta, \lambda) = f^0(\theta) + f'(\lambda) + if''(\lambda) \quad (2)$$

$\theta$  = scattering angle

$\lambda$  = wavelength

$f_0(\theta)$  = normal scattering term

$f''(\lambda)$  = absorption term

$f'(\lambda)$  = dispersive term

$i$  = imaginary unit

The last two terms represent the anomalous scattering that occurs at the absorption edge of the atom (see figure 10), whereby the absorption term is  $90^\circ$  advanced in phase. The atoms usually present in

proteins possess a negligible  $f'(\lambda)$  and  $f''(\lambda)$  due to the non-accessibility of their absorption edge by a X-ray beam. The K-absorption edge of S for example is at  $\lambda = 5.0155 \text{ \AA}$  and therefore not accessible at a synchrotron X-ray source. In contrast, the K-absorption edge of Se is at  $\lambda = 0.9795 \text{ \AA}$  and for this reason accessible (see figure 11). The substitution of S-atoms by exogenously provided Se within a protein causes the presence of atoms that scatter anomalously since their atomic form factor contains an imaginary contribution  $if''$  at the K-absorption edge (see figure 11). In this case Friedel's law is not valid anymore and both the amplitude and phase relationship of the Friedel pairs are broken. This is illustrated in figure 12A and becomes more evident when one Friedel mate is reflected across the horizontal axis (see figure 12B): The reflections  $F_{hkl}$  and  $F_{-h-k-l}$  exhibit different amplitudes and their phases are no longer complementary. The amplitude difference  $|F_{hkl}| - |F_{-h-k-l}|$  is known as the Bijvoet difference.

Anomalous scattering can be utilized for phasing. The missing phase values can only be obtained through some prior knowledge of the structure. One possibility to achieve such knowledge is the localization of anomalously scattering Se atoms in the electron density map of the molecule by solving the so called Patterson map.

The Patterson function is a variation of the Fourier sum (1) that utilizes the structure factors (amplitudes and phases) to obtain a map of the electron density. The Patterson function uses solely the measurable intensities (amplitudes squared) and is therefore independent from the phase values:

$$\rho(u, v, w) = \frac{1}{V} \sum_{hkl} |F_{hkl}|^2 e^{-2\pi i(hu + kv + lw)} \quad (3)$$

$\rho(u, v, w)$  = peak value in the Patterson map at position  $u, v, w$

$e$  = Euler's number

$V$  = volume of the unit cell

$i$  = imaginary unit

$|F_{hkl}|$  = structure factor amplitude of the reflection  $hkl$

The graph of this function is not a contour map of the electron density but rather a map that contains all interatomic vectors of the unit cell. The coordinates  $u, v, w$  locate a point in the Patterson map in the same way as coordinates  $x, y, z$  locate a point in the electron density map. Three peaks of three atoms in the electron density map give rise to 6 peaks in the Patterson map defined by moving the tails of the six possible interatomic vectors to the origin. The peaks of the Patterson map occur at the head of each vector. The origin and two of the six peaks in the map determine the location of the three real atoms in the unit cell. To solve the Patterson map, the origin and two additional peaks are chosen as a trial solution and the expected Patterson map for this specific arrangement of atoms is generated. The picked arrangement of the atoms is correct in case of congruence of the generated and the calculated Patterson map. For a small number of atoms (up to approximately 1000) it is possible to work out the original position of the

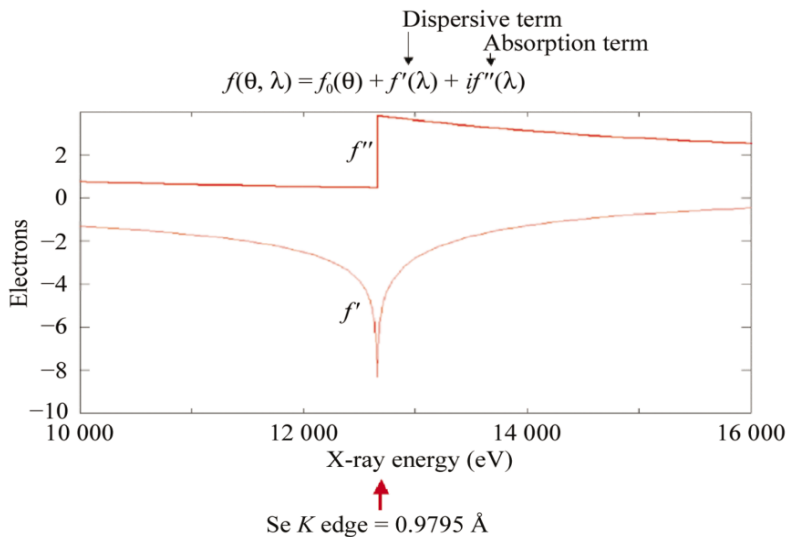
atoms in the unit cell by the described trial-and-error-method. Higher numbers of atoms ( $N$  real atoms in the electron density map give rise to  $N^2-N$  peaks in the Patterson map) make it difficult to solve a structure by deconvolving the Patterson map due to its increasing complexity. Even at high resolution, electron density peaks are barely resolved and hence interdict the construction of a Patterson map with peaks resolved from each other.

The Patterson map of a reflection dataset derived from a Se-methionine substituted protein crystal exhibits prominent peaks that correspond to the Se atoms in the unit cell. The height of a peak in the Patterson map is proportional to the product of the heights of the two corresponding peaks in the electron density map. Selenium has 34 electrons and sulfur 16 electrons. According to that, a selenium peak in the Patterson map exhibits almost the fivefold intensity of a sulfur peak. The intensity discrepancy between selenium and other elements usually present in proteins is even higher causing their fading into the noise. This allows the calculation of a Patterson map that exclusively encompasses the Se atoms within a molecule and is for this reason of low complexity. The direct solution of this Patterson map gives rise to the substructure of Se atoms in the electron density of the molecule. As soon as the locations  $x, y$  and  $z$  of the anomalous scattering atoms is known the corresponding phases can be calculated.

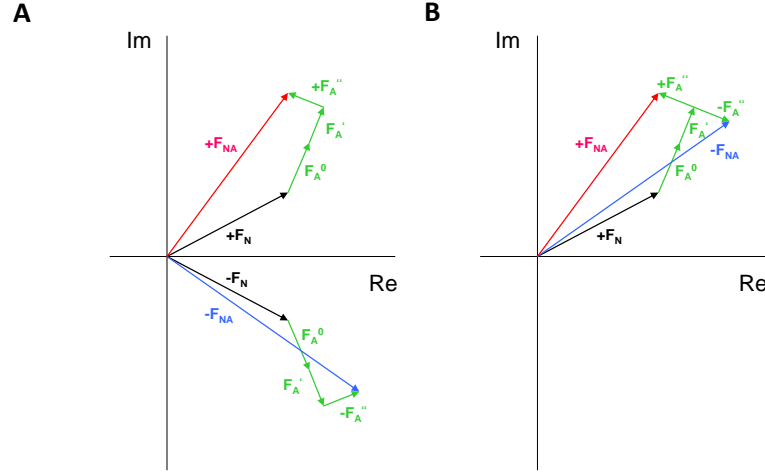
At this stage of prior knowledge of the anomalous substructure, the green vectors in figure 11 are known in length and phase and can be used to calculate the phase of the vector  $F_N$  and with it the phase of the normal scattering molecules in the molecule. In case of a SAD experiment two solutions are possible for the phase  $\alpha_N$  as illustrated in figure 13. Figure 13A shows a so called Harker construction in which the intersection points of the two circles define the two possible phases  $\alpha_N$  (highlighted in orange). Figure 13B represents an alternative way to illustrate the problem of the phase ambiguity. The two possible values of  $\alpha_N$  are symmetrically orientated about  $\alpha_A - 90^\circ$  and give rise to the same Bijvoet difference  $|+F_{NA}| - |-F_{NA}|$ . The unique determination of the protein phase is only possible when the phases  $\alpha_N$  and  $\alpha_A$  differ by  $90^\circ$ . In this case both solutions coincidence.

Different ways can be chosen to overcome the described phase ambiguity. One possibility is to carry out a MAD experiment (multiple anomalous dispersion). A SAD experiment is characterized by the measurement of anomalous scattering at one sole wavelength (usually at the absorption peak / K-edge ( $\lambda_1$ ) where the absorptive term  $F_A''$  exhibits a local maximum; see figure 13A). In a MAD experiment anomalous scattering is recorded at least at one more wavelength (inflection point ( $\lambda_2$ ) where the dispersive term  $F_A'$  exhibits a local minimum, remote ( $\lambda_3, \lambda_4$ ); see figure 14A). This procedure enables the determination of a further Bijvoet difference (see figure 14B) that allows complementation of the Harker construction shown in figure 12A by a third circle. This results in one coincident intersection point of all three circles, which defines a unique solution of  $\alpha_N$ .

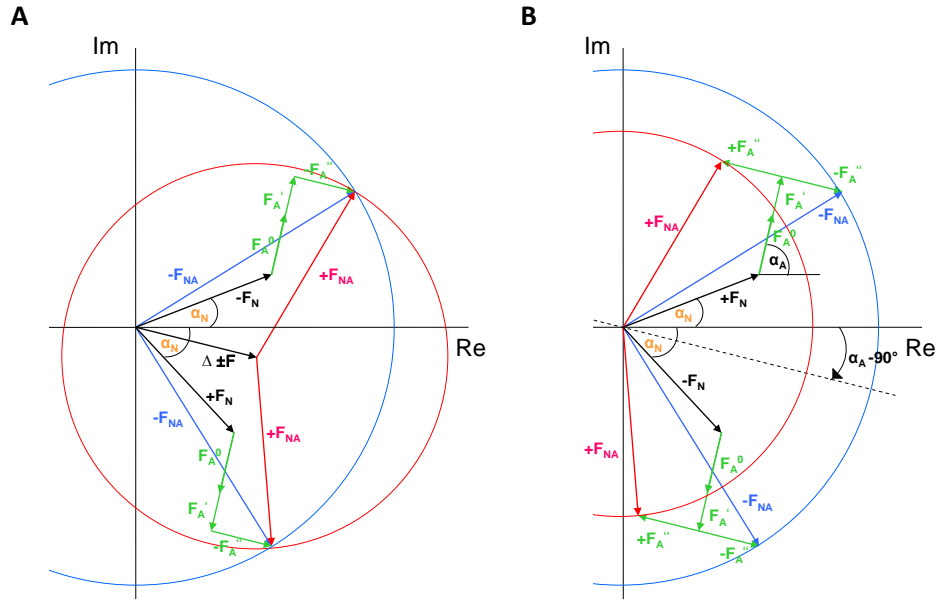
Another possibility is to collect data at just one single wavelength and to utilize density-modification protocols to break the phase ambiguity. One method of density modification is solvent flattening that sets the electron density in the solvent regions to a typical value differing from the typical electron density of protein. This procedure smoothes the electron density map by eliminating random fluctuations in density and divides the map into two main subregions: those of relative high density caused by protein and those of relative low density caused by solvent. Flattening of the solvent region leads to higher accuracy of the corresponding phases and an enhanced contrast between protein and solvent. Solvent flattening requires phases, which are at least good enough to allow a boundary definition between the disordered solvent and the ordered protein. The solvent-protein boundary can be defined automatically by utilizing different algorithms.



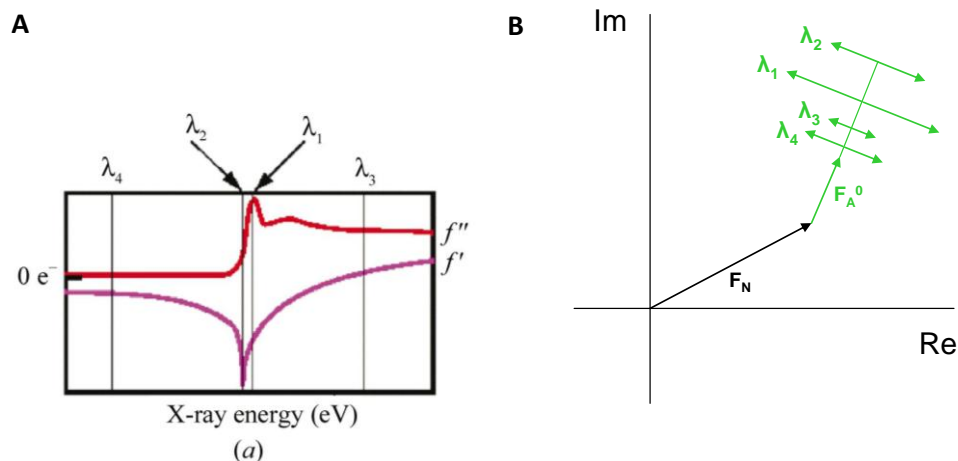
**Figure 11:** Variation in anomalous scattering of Se plotted against the energy of the X-ray beam (145).



**Figure 12:** (A) Breakdown of Friedel's law in presence of an anomalous scatterer. The total structure factors  $F_{NA}$  of the Friedel mates are shown in red and blue respectively. The contribution of normally scattering atoms is shown in black ( $F_N$ ) and that of anomalous scatterers in green ( $F_A^0$ ,  $F_A'$ ,  $F_A''$ ). The total structure factors  $F_{NA}$  contain the normal scattering of all atoms  $F_N$  (shown in black) and  $F_A^0$  (shown in green), the dispersive component of the anomalously scattering atoms  $F_A'$  and the absorptive component of the anomalously scattering atoms  $F_A''$ . (B) Reflection of one Friedel mate across the horizontal axis illustrates the invalidity of Friedel's law in the presence of an anomalous scatterer:  $|F_{hkl}| \neq |F_{-h-k-l}|$ ;  $\alpha_{hkl} \neq -\alpha_{-h-k-l}$  (adapted from (144)).



**Figure 13:** SAD experiments result in a phase ambiguity which can be illustrated by a Harker construction (A) and an Argand diagram (B) showing the various contributions to the scattering factors. The total structure factors of the Friedel  $F_{NA}$  mates are shown in red and blue respectively. The contribution of normally scattering atoms is shown in black ( $F_N$ ) and that of the anomalous scatterers in green ( $F_A^0$ ,  $F_A'$ ,  $F_A''$ ) (adapted from (146)).



**Figure 14:** MAD phasing. (A) Typical absorption curve of an anomalous scatterer (145). (B) Measurement of anomalous scattering at more than one wavelength allows the determination of further Bijvoet differences and breaks the phase ambiguity (adapted from (144)).

#### 2.2.12.3 Structure determination of the ssMDA5:PIV5 V protein core complex

The structure of the ssMDA5:PIV5 V protein core complex was determined by SAD phasing. Diffraction data were collected at beamline ID23-2 of the European Synchrotron Radiation Facility (ESRF). Diffraction images of selenomethionine-substituted protein crystals were recorded at a wavelength of 12.66127 keV (corresponds to the Se peak wavelength) with an oscillation range of  $\Delta \varphi = 1^\circ$  per image. Processing of the collected diffraction data was carried out with XDS (147). The crystals contained one complex per asymmetric unit (37% solvent content) and belong to the space group  $P2_12_12_1$  with cell constants  $a = 51.07 \text{ \AA}$ ,  $b = 52.49 \text{ \AA}$ ,  $c = 120.79 \text{ \AA}$ ,  $\alpha = \beta = \gamma = 90^\circ$ . The integrated and scaled reflection intensities/amplitude listed in the XDS output file allowed processing up to  $2.3 \text{ \AA}$ . The structure was determined by single wavelength anomalous dispersion using SHELXD (as part of AutoSHARP (148)) to locate selenium sites. Phase improvement was achieved by solvent-flattening using SOLOMON (as part of AutoSHARP (148)). The optimized experimental phases resulted in an electron density map whose quality was sufficient to operate initial automatic model building with Buccaneer (149). Besides of 5 anomalous peaks in the anomalous difference map representing expected selenium atoms from the selenomethionine-substituted ssMDA5 protein, two additional anomalous peaks of zinc atoms belonging to the PIV 5 V protein could be identified. The initial model was refined in cyclic rounds of manual model building in COOT (150) and refinement using PHENIX (151). Refinement represents an iterative process in which the atomic model undergoes modification in usually five parameters for each atom: the three coordinates  $x, y$  and  $z$  which specify the position of each atom within the real space of the unit cell,

the B-factor (also known as temperature factor) which describes the attenuation of elastic x-ray scattering depending on the temperature and the occupancy Q. Occupancy Q represents the time that an atom spends at position  $x,y,z$ . Based on the modified model, structure factor amplitudes are calculated and the discrepancy between the calculated amplitudes  $|F_{\text{calc}}|$  and the observed amplitudes  $|F_{\text{obs}}|$  is determined. The goal is to achieve a model that exhibits maximal agreement between  $|F_{\text{calc}}|$  and  $|F_{\text{obs}}|$ . A refinement process can therefore be seen as a least squares method of the target function  $\sum (|F_{\text{obs}}| - k|F_{\text{calc}}|)^2$ , whereby k is a scale factor necessary to bring  $|F_{\text{calc}}|$  and  $|F_{\text{obs}}|$  on the same scale. The final model comprising ssMDA5 residues 546-808 and PIV5 V protein residues 168-219 showed good R-factors and stereochemistry. Atomic coordinates and structure factors were deposited in the Protein Data Bank with accession number 4IIS. The crystallographic parameters of the structure are listed in tables 4 and 5 (see section 3.4). All figures were prepared with PyMol.

### 2.2.13 Analytical size exclusion chromatography

Purified mmMDA5 protein or mmRIG-I protein (purified as described in (55)) was diluted to a concentration of 15  $\mu\text{M}$ . PIV5 V protein was added in molar excess (1:1.5 ratio) and the complex was incubated for 1h on ice. Analytical size exclusion experiments were performed on an Ettan LC system (Superdex200 5/150, GE Healthcare) using the complex formation buffer (see table 3) as running buffer.

### 2.2.14 Differential scanning fluorimetry experiments

Differential scanning fluorimetry experiments were carried out with 1 $\mu\text{g}/\mu\text{l}$  dilutions of MDA5 full length or MDA5 SF2 ATPase domain with or without PIV5 V-protein (purified stoichiometric complex) in 50 mM Tris pH 7.5, 150 mM NaCl, 5 mM  $\text{MgCl}_2$ , 10  $\mu\text{M}$   $\text{ZnCl}_2$ , 5 mM DTT and a 5 $\times$  dilution of SYPRO Orange dye (Invitrogen) as described (152). The fluorescence signal as a function of temperature was recorded using a Real Time PCR machine (Bio-Rad). The temperature gradient was performed in the range of 10–70  $^{\circ}\text{C}$  with a ramp of 1  $^{\circ}\text{C}$  over the course of 60 min.

### 2.2.15 Electrophoretic mobility shift assays (EMSAs)

A 24bp RNA-palindrome purchased from Biomers (HPLC-grade) with following sequence was used: 5'-GCGUCGUACGCUAGCGUACGACGC-3'. A fixed RNA concentration of 1  $\mu\text{M}$  was incubated with the indicated amount of purified mmMDA5 in 50 mM Tris pH 8.0, 150 mM NaCl, 5 mM  $\text{MgCl}_2$ , 10  $\mu\text{M}$   $\text{ZnCl}_2$  and 5 mM DTT reaction buffer. PIV5 V protein was added in 3-fold molar excess according to the



MDA5 concentration. Each incubation step was carried out for 40 min on ice. Samples were separated by native PAGE (5%) and analyzed by Gel-Red-staining (Biotium).

### **2.2.16 Electron microscopy**

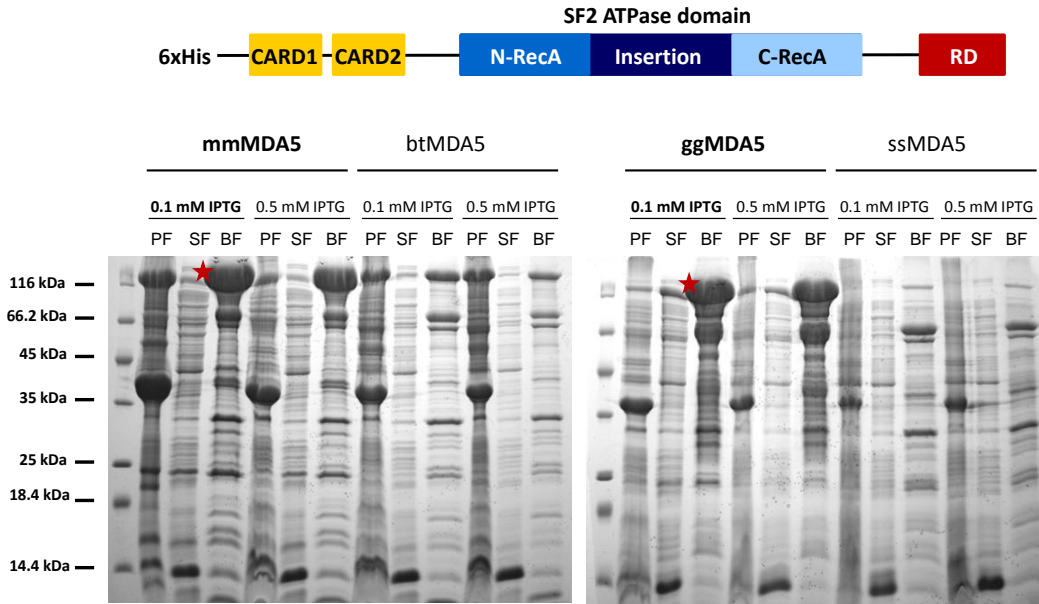
A solution containing 125 µg/ml of poly(I:C) (The Midland Certified Reagent Company) and 50 µg/ml mmMDA5 were incubated in absence or presence of 2.5 mM ATP and PIV5 V protein (in 1.2 molar excess to mmMDA5). The incubation was carried out for 40 min on ice at pH 7.4 in 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 50 mM Glycine, 14 mM Succinic acid, 10 mM MgCl<sub>2</sub>, 10 µM ZnCl<sub>2</sub> and 5 mM DTT. The mixtures were applied to pre-coated Quantifoil holey carbon-supported grids and negatively stained using 2% uranyl acetate (performed by Charlotte Ungewickell, laboratory of Prof. Dr. Roland Beckmann, Gene Center, Munich). Grids were examined with the help of Günter Pfeifer (Max Planck Institute of Biochemistry, Martinsried) on a CM20 FEG Philips at 160 kV and images were recorded with a 4k x 4k CCD camera (Tietz).

### 3. Results

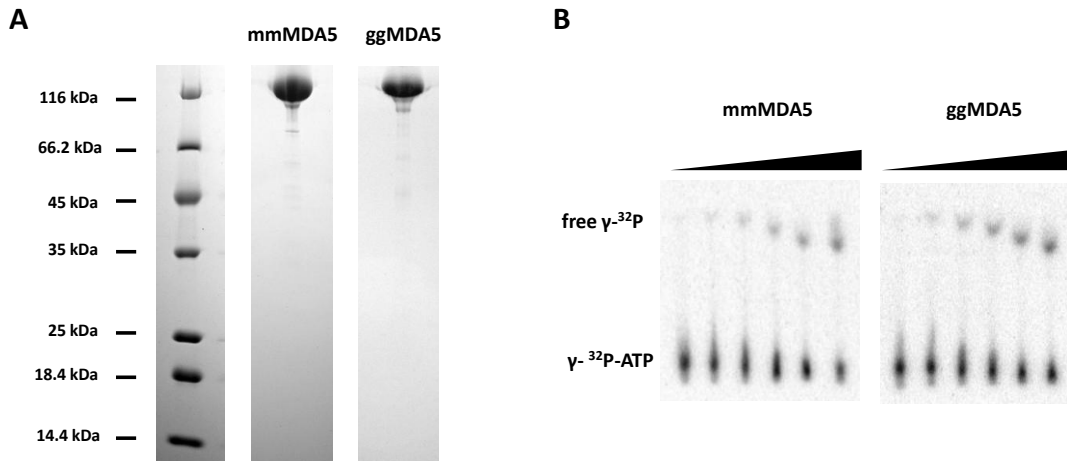
#### 3.1 Recombinantly expressed full length MDA5 is a functional protein

At the beginning of the work presented in this thesis, MDA5 was poorly characterized in terms of its atomic structure. With the exception of the isolated MDA5 RD, no structural information concerning further domains or their interplay in viral RNA ligand recognition existed. Crystallization of MDA5 and other RLRs was a major goal as the key towards understanding of their detailed function lies in the molecular structure of the receptors. To achieve the task of MDA5 crystallization in complex with the inhibitory paramyxoviral V protein and to gain maximal structural insight into the receptor, different full length MDA5 constructs of different species were cloned and tested for soluble expression in *E.coli*. Mouse (mm, *mus musculus*) and chicken (gg, *gallus gallus*) MDA5 recombinantly expressed as His-tagged versions were found to be soluble and to be produced in high yields required for crystallization (see figure 15). Both full length proteins could be purified to sufficient homogeneity (see figure 16A) and showed no sign of aggregation in size exclusion chromatography (see figure 17). Activity assays revealed that mmMDA5 as well as ggMDA5 purified from *E.coli* were able to hydrolyse ATP in the presence of the synthetic dsRNA ligand poly(I:C) (see section 1.2.1), indicating that both proteins are natively folded and functional (see figure 16B).

The two purified MDA5 orthologs were tested for their ability to bind purified PIV5 V protein. Both, mmMDA5 as well as ggMDA5, were found to form a stoichiometric complex with the V protein that was stable in size exclusion chromatography (see figure 17). Further investigation using static light scattering (SLS) measurements (see section 2.2.7) revealed that MDA5 is a monomeric protein in solution and binds the PIV5 V protein with a stoichiometry of 1:1 (see figure 18). Complexes comprising full length MDA5 and full length PIV5 V protein isolated by size exclusion chromatography were of sufficient quality (see figure 17B) for crystallization trials. However, none of the purified full length complexes tested under various screening conditions was observed to form crystals.

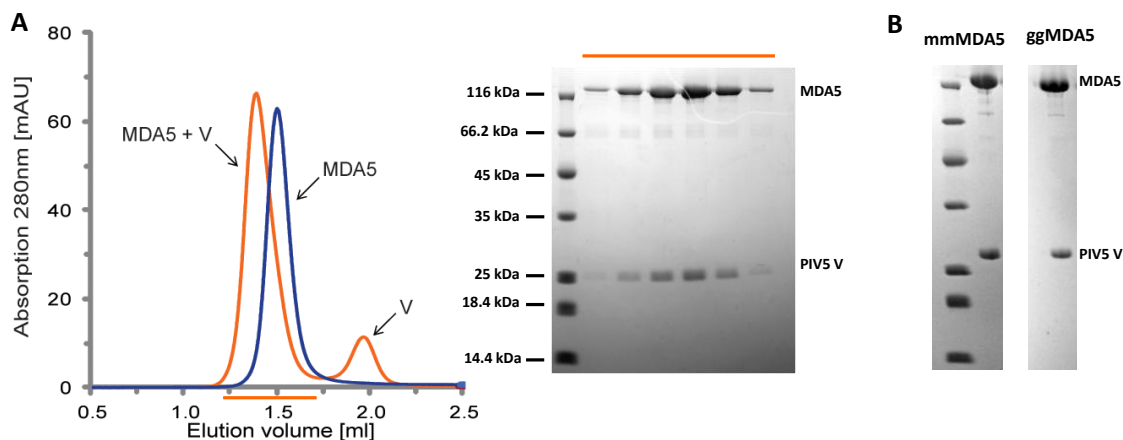


**Figure 15:** Testexpression of N-terminally His-tagged full length MDA5 of different orthologs (mm = *mus musculus*, bt = *bos taurus*, gg = *gallus gallus*, ss = *sus scrofa*) analyzed by SDS-PAGE and Coomassie staining. Recombinant expression was performed in *Escherichia coli* Rosetta (DE3) for 18h at 20°C with indicated ITPG concentrations. PF = pellet fraction containing unsoluble protein. SF = soluble protein fraction. BF = bead fraction (soluble proteins eluted from Ni-NTA agarose beads). Conditions and constructs that gave rise to best expression yields of soluble protein are highlighted in bold and indicated by red stars.

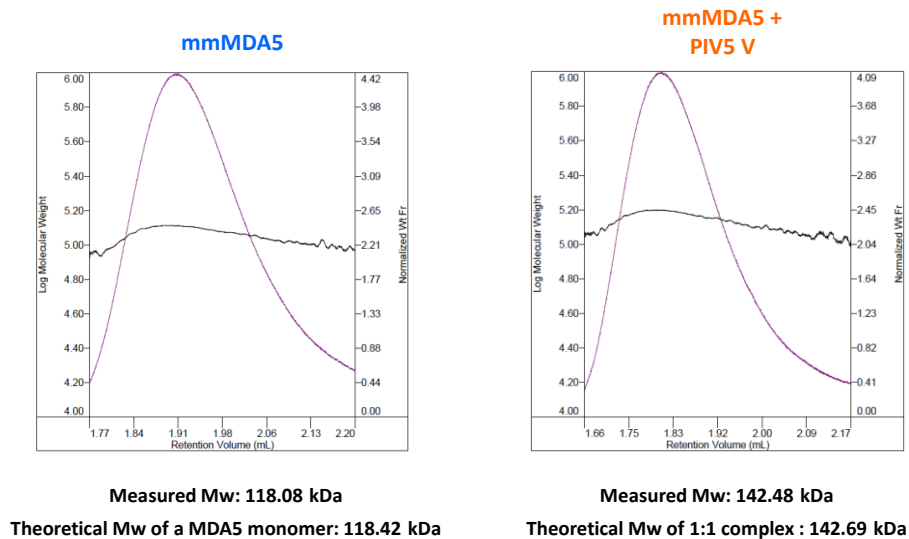


**Figure 16:** (A) Purified full length mmMDA5 and ggMDA5 analyzed by SDS-PAGE and Coomassie staining. (B) ATPase assay showing the ATPase activity of the purified full length proteins used in crystallization trials and biochemical assays. Poly(I:C)-stimulated mmMDA5 or ggMDA5 (0  $\mu$ M, 1  $\mu$ M, 2.5  $\mu$ M, 5  $\mu$ M, 10  $\mu$ M and 15  $\mu$ M) was incubated with  $\gamma$ - $^{32}$ P-ATP. Liberated and TLC-separated  $\gamma$ - $^{32}$ P was visualized with a phosphorimager.

### III RESULTS



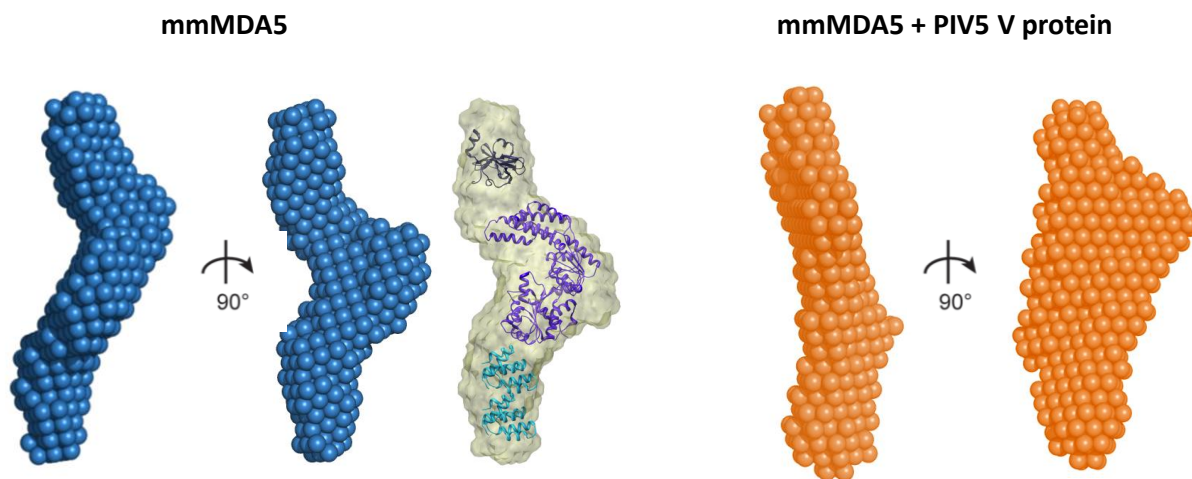
**Figure 17:** (A) Analytical size exclusion chromatogram of full length mmMDA5 in absence (blue elution profile) and presence of purified PIV 5 V protein (orange elution profile). V protein binding is reflected in an increase of the hydrodynamic radius causing a shift of the protein peak to smaller elution volume. The solid line in orange marks fractions analyzed by SDS-PAGE and Coomassie staining revealing stoichiometric coelution of full length mmMDA5 and PIV5 V protein. (B) SDS-PAGE analysis (Coomassie staining) of finally purified and concentrated full length MDA5:PIV5 V protein complexes used in crystallization trials.



**Figure 18:** Multiangle static light scattering experiments in combination with size exclusion chromatography were performed to determine the absolute molar mass of full length mmMDA5 in isolation (left panel) and in complex with PIV5 V protein (right panel) in solution. The measured mean values of the molecular weight (Mw) of isocratically eluted mmMDA5 and mmMDA5:PIV5 V protein complex testify a monomeric state of full length MDA5 and a 1:1 stoichiometry of the MDA5:PIV5 V protein complex in solution (theoretical Mw of the PIV5 V protein: 24.27 kDa).

### 3.2 Optimization of protein constructs for crystallization

*Ab initio* protein models calculated from SAXS data revealed that monomeric full length MDA5 adopts an elongated, non-globular conformation in solution (see figure 19). The shape of the MDA5 full length molecule is characterized by several lobes and defined by multiple centers of mass (see also section 2.2.8 and figure 33 of section 3.6). The extended, multipart shape of the protein suggests that the single domains (CARDs, SF2 ATPase domain, RD) of free MDA5 might be connected by flexible linkers allowing them to change their relative position to each other within the MDA5 molecule. Suchlike linkage would explain why full length MDA5 is prone to degradation in testexpression as well as in purification (see figures of section 3.1) indicating that non-compact, flexible parts are exposed and accessible for proteases. Binding of the PIV5 V protein to MDA5 was not found to cause a compaction of the MDA5 full length molecule. *Ab initio* models calculated from SAXS data derived from purified MDA5:PIV 5 V protein complexes showed that V protein binding rather enhances elongation of the MDA5 molecule that appears to be stretched by the V protein (see figure 19 and also figure 33 of section 3.6).

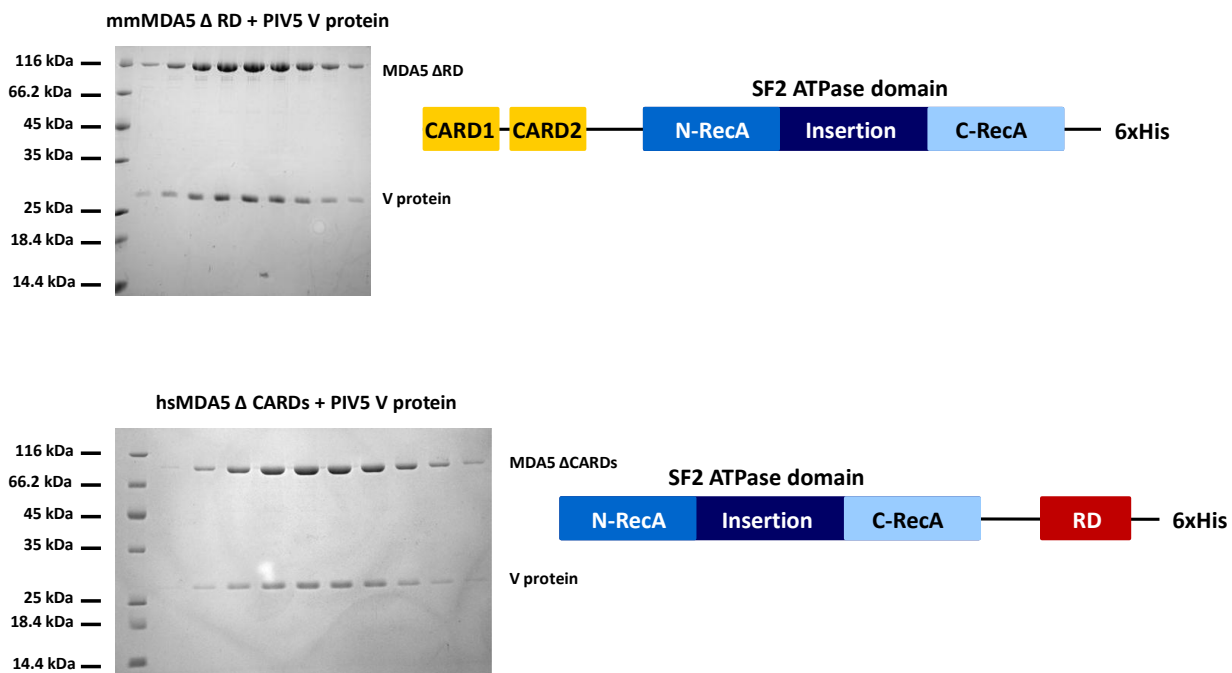


**Figure 19:** *Ab initio* protein models of mmMDA5 alone (shown in blue) and in complex with PIV5 V protein (shown in orange) obtained by SAXS analysis. The Apaf 1 CARD (shown in cyan, PDB code: 2P1H), the Hef SF2 ATPase domain (shown in mauve, PDB code: 1WP9) and the RD of hsMDA5 (shown in black, PDB code: 3GA3) were modeled into the electron density shape of full length mmMDA5.

The assumed intrinsic flexibility of the MDA5 molecule was considered to prevent or hinder crystallization of the full length MDA5:PIV5 V protein complex. To address this problem, further MDA5 constructs combining different domains of the protein were purified in complex with PIV5 V protein and

### III RESULTS

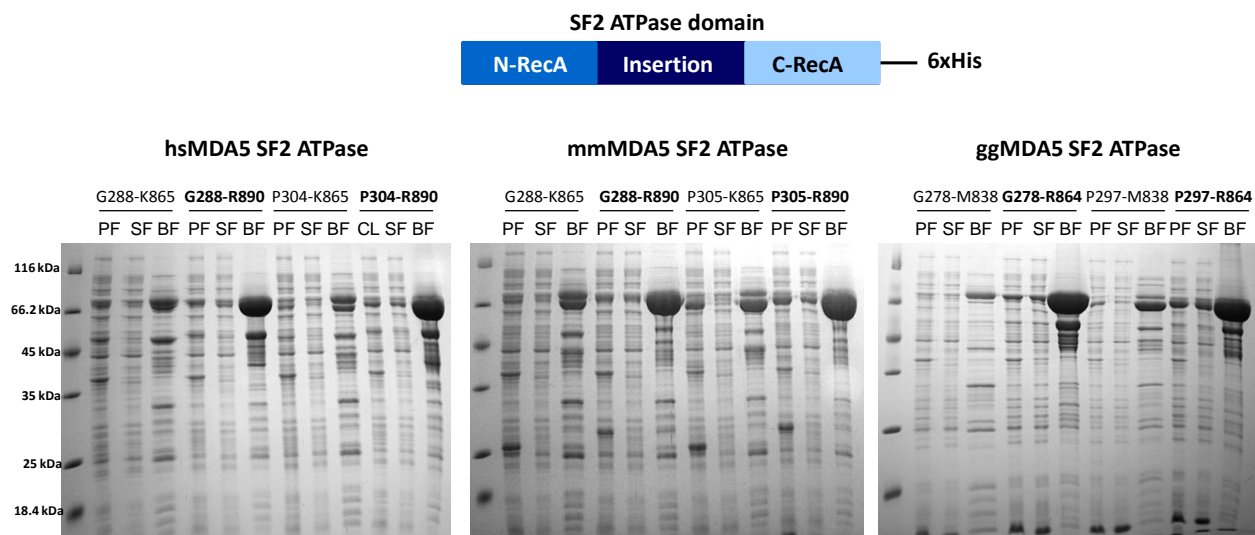
screened for crystallization. Figure 20 shows representative complexes of PIV5 V protein and RD- or CARD-less MDA5 constructs after size exclusion chromatography. None of the truncated complexes tested under various screening conditions was observed to form crystals.



**Figure 20:** SDS-PAGE analysis (Coomassie staining) of truncated MDA5:PIV5 V protein complexes purified by size exclusion chromatography and used in crystallization trials. The C-terminally His-tagged MDA5 constructs are illustrated in schemes. Hs = *homo sapiens*, mm = *mus musculus*.

In a next step, as RD- and CARD-less MDA5 constructs did not succeed in crystallization, various constructs of the MDA5 SF2 ATPase domain with different N- and C-termini were generated and tested for soluble expression in *E.coli*. Figure 21 shows the testexpression results of C-terminally His-tagged SF2 ATPase versions of three different MDA5 orthologs. Constructs with a C-terminus solely comprising the C-terminal RecA-like fold of the SF2 ATPase domain were found to be either unexpressed or low expressed. In contrast, constructs whose C-terminus contain the C-terminal RecA-like fold and two downstream, well conserved  $\alpha$ -helices (according to amino acid sequence alignments and secondary structure predictions) gave rise to high expression levels (see figure 21). These two additional  $\alpha$ -helices that characterize the SF2 ATPase domain of all RLR family members were later found to form a functionally important, V-shaped linker within the RNA clamp structure of RIG-I and MDA5 (see sections 1.3.1 and 1.3.2).

### III RESULTS

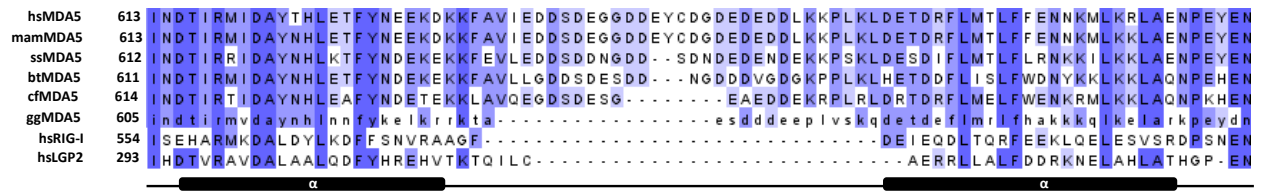


**Figure 21:** Testexpression of C-terminally His-tagged MDA5 SF2 ATPase domain constructs (N-and C-termini as indicated) of different orthologs (hs = *homo sapiens*, mm = *mus musculus*, gg = *gallus gallus*) analyzed by SDS-PAGE and Coomassie staining. Recombinant expression was performed in *Escherichia coli* Rosetta (DE3) for 18h at 20°C with final IPTG concentration of 0.1 mM IPTG. PF = pellet fraction containing unsoluble protein. SF = soluble protein fraction. BF = bead fraction (soluble proteins eluted from Ni-NTA agarose beads). Constructs that gave rise to expression yields sufficient for crystallization are highlighted in bold.

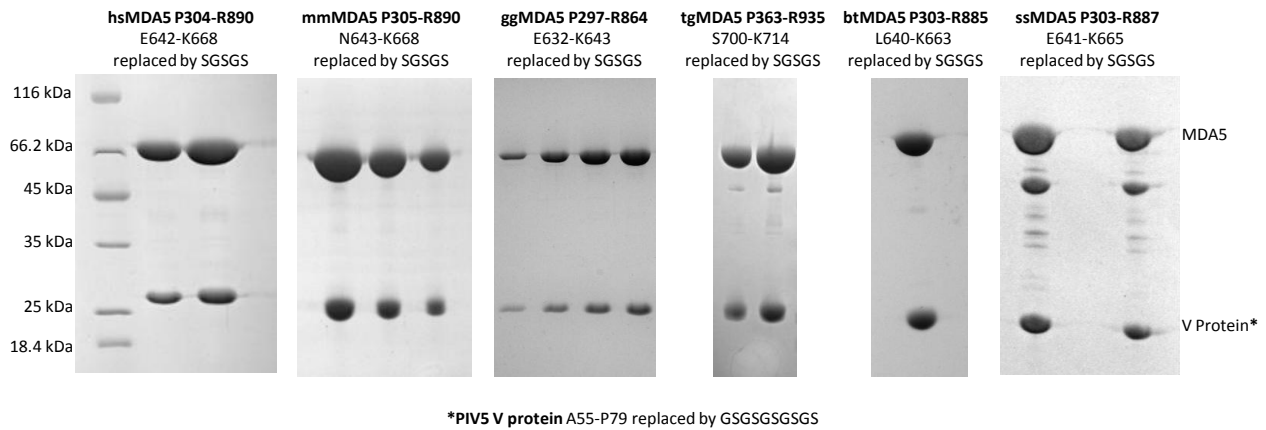
SF2 ATPase domain constructs proven to be expressed in sufficient yields (see figure 21) were purified in complex with PIV5 V protein and screened for crystallization. None of the SF2 ATPase:PIV5 V protein complexes tested under various screening conditions was observed to form crystals.

Amino acid sequence alignments of RLR family members revealed a loop region within the  $\alpha$ -helical insertion of the MDA5 SF2 ATPase domain, that is missing in case of LGP2 as well as RIG-I and predicted to be unstructured (see figure 22). To optimize MDA5 SF2 ATPase constructs for crystallization, this likely flexible loop was replaced by a short SGSGS-linker. A similar strategy was applied in case of the PIV5 V protein. A large loop region of 25 amino acids found to be unstructured in this V protein (104) was replaced by a GSGSGSGSGS-linker to increase the probability of crystallization. Figure 23 shows engineered MDA5 proteins of six MDA5 orthologs after purification and complex formation with modified PIV5 V protein. None of the optimized SF2 ATPase:PIV5 protein complexes tested under various screening conditions was observed to form crystals.

### III RESULTS



**Figure 22:** Amino acid sequence alignment of the SF2 ATPase loop region of selected RLR family members with annotated secondary structure elements. Residue conservation is color-coded according to percentage identity. Hs = *homo sapiens*, mam = *macaca mulatta*, ss = *sus scrofa*, bt = *bos taurus*, cf = *canis familiaris*, gg = *gallus gallus*.



**Figure 23:** SDS-PAGE analysis (Coomassie staining) of finally purified and concentrated MDA5 SF2 ATPase:PIV5 V protein complexes with flexible loops replaced by GS-linkers as indicated and used in crystallization trials. Hs = *homo sapiens*, mm = *mus musculus*, gg = *gallus gallus*, tg = *taeniopygia guttata*, bt = *bos taurus*, ss = *sus scrofa*.

### 3.3 Crystallization of the ssMDA5:PIV5 V protein core complexes

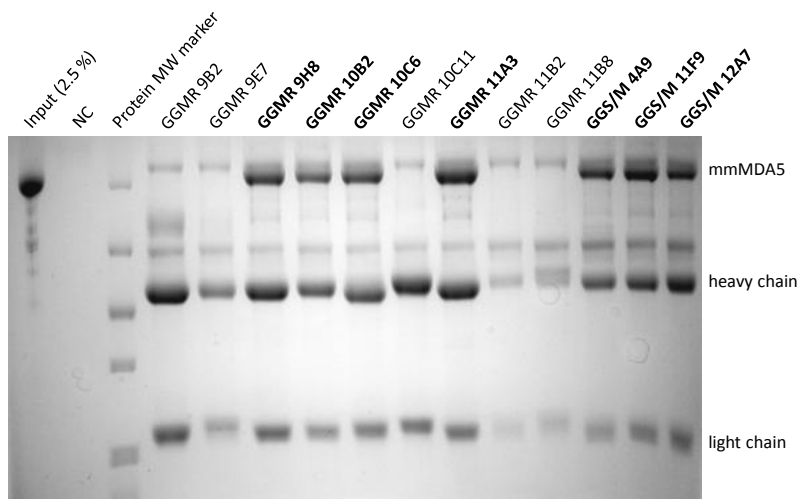
One major issue regarding crystallization of the MDA5:V protein complex is its strong tendency to form heavy precipitate in crystallization screens. Even at very low protein concentrations (in the range of 0.5-1.0 mg/ml) each and all of the described protein constructs immediately precipitated in crystallization experiments. Crystals are formed when proteins “precipitate” very slowly from a supersaturated solution in an ordered manner. The misbehavior of the MDA5:V protein complex made it impossible to define a reasonable protein concentration range that would allow such a crystallization process and to identify useful conditions for crystal generation. Alteration of major parameters known to influence crystallization and crystal growth (e.g. temperature, nature and concentration of the precipitant, organic and inorganic



### III RESULTS

additives) did not resolve the problem of strong precipitation. Reductively methylated MDA5 proteins purified in complex with PIV5 V protein and complexes purified in high salt buffers (the MDA5:PIV5 V protein complex was found to be stable at 1M NaCl in size exclusion chromatography) did not increase solubility either.

Another strategy that was applied to prevent immediate precipitation of the MDA5:V protein complex in screening experiments was its co-crystallization with Fab fragments. It was further assumed that Fab fragments might stabilize the MDA5:PIV5 V protein complex and allow crystallization by providing crystal contacts. The Fab fragments were derived from monoclonal antibodies recognizing either mmMDA5 or PIV5 V protein (see section 2.2.10). Figure 24 shows a pulldown assay that was performed with supernatants of different hybridoma cell lines to identify monoclonal antibodies that were able to bind natively folded full length mmMDA5 purified from *E.coli*. Further selection rounds based on similar pulldown approaches were used to identify PIV 5 V protein binders, binders that specifically recognize the MDA5 SF2 ATPase domain and binders that do not interfere with complex formation of the two proteins.

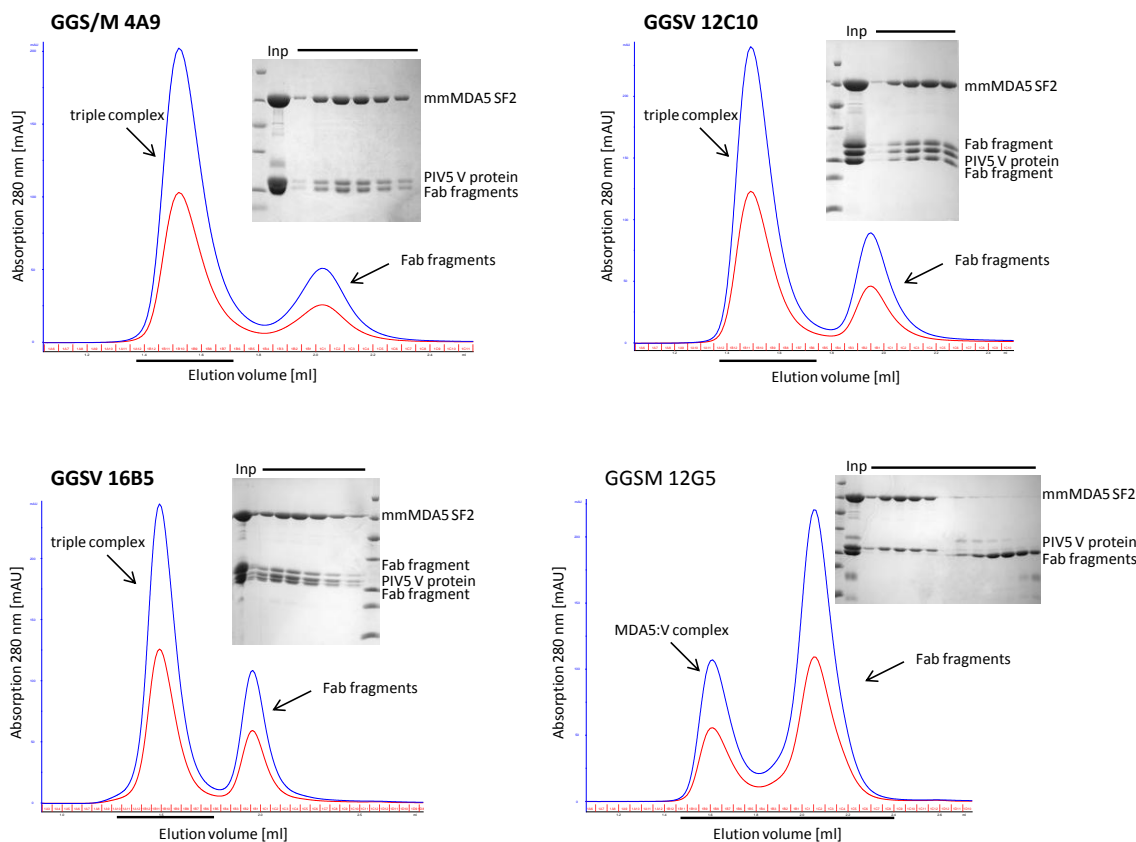


**Figure 24:** Pulldown assay applied to identify monoclonal IgG antibodies that are able to recognize natively folded full length mmMDA5 purified from *E.coli* and analyzed by SDS-PAGE and Coomassie staining. The assay was performed with supernatants of indicated hybridoma cell clones and protein G beads. NC = negative control performed with full length mmMDA5 and protein G beads. Positive clones are highlighted in bold.

IgG antibodies of four positive clones (namely the mmMDA5 recognizing clones GGS/M 4A9, GGS/M 12G5 and the PIV5 V protein recognizing clones GGSV 12C10, GGSV 16B5) were subjected to papain digest. The produced and purified Fab fragments were tested for their ability to form stable,

### III RESULTS

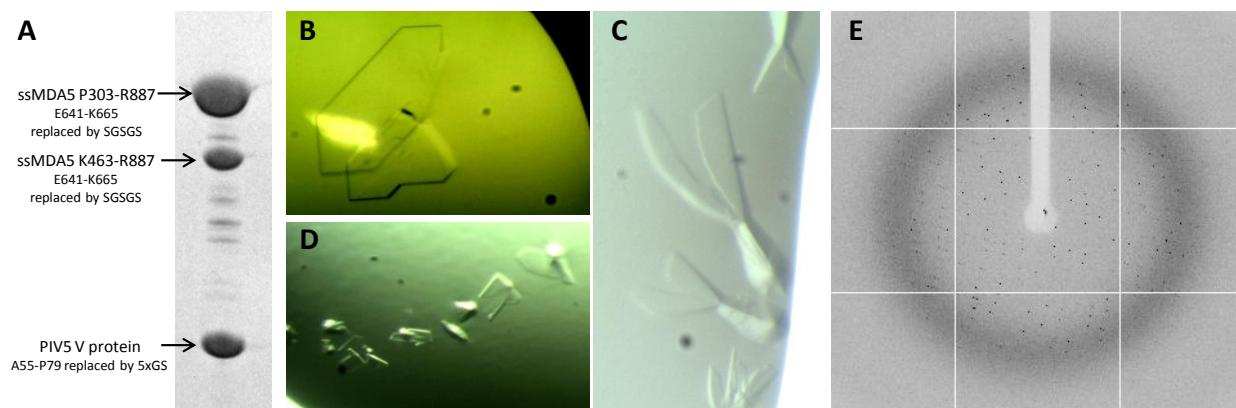
stoichiometric triple complexes with preformed MDA5:V protein complexes in size exclusion chromatography (see figure 25). Three high affinity binders (namely GGS/M 4A9, GGSV 12C10 and GGSV 16B5) could be identified and were used for co-crystallization trials with different MDA5:PIV5 V protein constructs. Although binding of Fab fragments significantly increased the solubility of the MDA5:PIV5 V protein complex in crystallization experiments, none of the Fab triple complexes tested under various screening conditions was observed to form crystals.



**Figure 25:** Analytical size exclusion chromatography experiments performed to select Fab fragments of sufficient affinity to form stable triple complexes with preformed MDA5:PIV5 V protein complexes. Positively tested Fab fragments are highlighted in bold. Blue elution profiles are based on the absorption at 280 nm, red elution profiles are based on the absorption at 260 nm. MmMDA5 SF2 = mus musculus MDA5 SF2 ATPase domain. Solid lines mark fractions analyzed by SDS-PAGE and Coomassie staining.

The goal of MDA5:PIV5 V protein crystallization was finally achieved by an in-drop proteolysis approach. The loop-truncated complexes of MDA5 SF2 ATPase domains and PIV5 V protein shown in figure 23 were mixed with different proteases (subtilisin, chymotrypsin or trypsin) before being used in

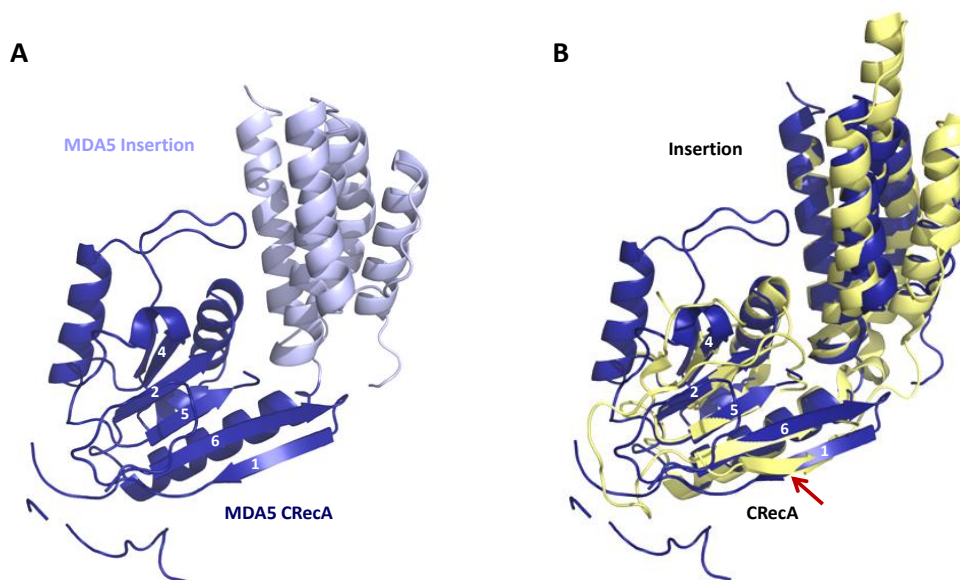
crystallization experiments. Among six complexes tested, the truncated ssMDA5 SF2 ATPase domain in complex with truncated PIV5 V protein (see figure 26A) was found to form crystals in the presence of trace amounts of trypsin (see figure 26B and section 2.2.11). The plate-shaped crystals obtained in commercial crystallization screens were optimized and converted to a more three-dimensional shape by hanging-drop vapour diffusion and streak seeding (see figures 26C and 26D). Native as well as selenomethionine-substituted protein crystals diffracted to a resolution of 2.3 Å (see figure 26E).



**Figure 26:** (A) Purified MDA5:PIV5 V protein complex that crystallized in the presence of trypsin. The co-purified impurity could be identified as N-terminally truncated fragment of the ssMDA5 SF2 ATPase domain. (B) Initial plate-shaped crystals that appeared in a commercial crystallization screen (condition: 100 mM Hepes pH 7.5, 20% (w/v) PEG 1500). (C) Refined crystals in the final condition (100 mM Hepes pH 7.5, 18% (w/v) PEG 1500). (D) Crystals optimized by streak seeding in the final condition. (E) Diffraction image of selenomethionine-substituted protein crystals recorded at beamline ID23-2 of the European Synchrotron Radiation Facility (ESRF).

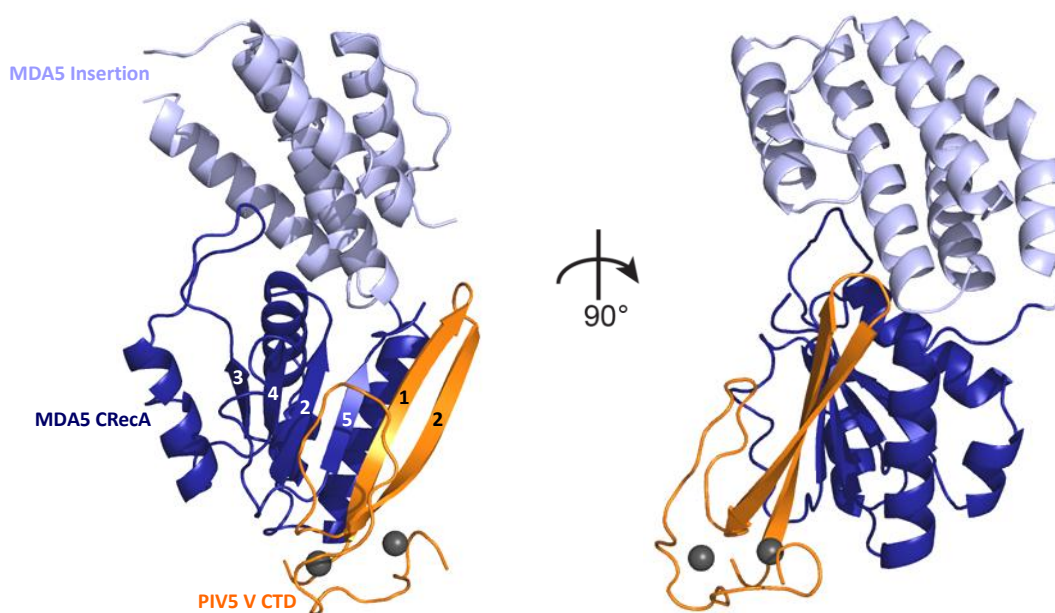
### 3.4 The paramyxoviral V protein disrupts the fold of the MDA5 SF2 ATPase domain

The presence of a protease in the crystallization condition raised the question which actual portion of the purified ssMDA5 SF2 ATPase:PIV5 V protein complex the crystals contain. The selenomethionine derivative of the ssMDA5 SF2 ATPase in complex with unlabeled PIV5 V protein crystallized in the same condition as the native complex and allowed structure determination by experimental phasing using SAD data (see section 2.2.12.3). The quality of the calculated electron density map was sufficient to place a first model (see figure 27A) using the autobuilding program Buccaneer (149). According to this initial model the crystals apparently contained the  $\alpha$ -helical insertion as well as the C-terminal RecA-like fold of the ssMDA5 SF2 ATPase domain, but no portion of the PIV5 V protein.



**Figure 27:** (A) Initial unrefined structure model obtained with Buccaneer (149). The C-terminal RecA-like fold of the ssMDA5 SF2 ATPase domain with annotated  $\beta$ -strands is shown in dark blue and the  $\alpha$ -helical insertion domain in light blue. (B) Superposition of the ssMDA5 model shown in (A) (in dark blue) with equivalent parts of the mmRIG-I SF2 ATPase domain (PDB code: 3TBK) (in yellow). The opposite orientation of  $\beta$ -strand 1 within the C-terminal RecA-like fold is indicated by a red arrow.

Superposition of the model obtained with Buccaneer and equivalent parts of the RIG-I SF2 ATPase domain confirmed their high similarity in fold and arrangement, but demonstrated opposite orientation of the  $\beta$ -strand 1 within the C-terminal RecA-like fold (see figure 27B). Refinement of the initial model by manual model building revealed further that residues of ssMDA5 belonging to the  $\beta$ -strands 1 and 6 within the C-terminal RecA-like fold did not fit in the assigned part of the electron density map. This paradox and the knowledge that the PIV5 V protein contains a zinc binding motif suggested the calculation of an anomalous difference Patterson map that uses the anomalous difference at the peak wavelength to locate anomalous scatterers. Besides of 5 anomalous peaks representing expected selenium atoms from the selenomethionine-substituted ssMDA5 protein, two additional anomalous peaks of zinc atoms could be identified. This information revealed the actual identity of the  $\beta$ -strands 1 and 6 assumed to be part of the ssMDA5 protein. These two  $\beta$ -strands do not belong to the C-terminal RecA-like fold of MDA5, but rather represent the conserved zinc binding motif of the V protein CTD (see section 1.5.1). Thus, the structure represents the evolutionary conserved, necessary and sufficient core of the MDA5:V protein complex (131, 134-136) (see also section 1.5.3). Figure 28 shows the final refined structure of the ssMDA5:PIV5 V protein core complex with the assigned zinc ions of the V protein CTD. Crystallographic data and refinement statistics are summarized in tables 4 and 5.



**Figure 28:** Structure of the ssMDA5:PIV5 V core complex with annotated  $\beta$ -strands. The ssMDA5  $\alpha$ -helical insertion is shown in light blue, the ssMDA5 C-terminal RecA-like fold in dark blue and the zinc binding motif of the PIV5 V C-terminal domain (CTD) in orange. Zinc ions are highlighted in grey.

**Table 4:** Summary of crystallographic data (\*highest resolution shell is shown in parenthesis).

Data collection	
Beamline	ESRF ID23-1
Wavelength (Å)	0.9792
Spacegroup	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	51.07, 52.49, 120.79
$\alpha = \beta = \gamma$ (°)	90
Resolution (Å)*	50-2.3 (2.35-2.29)
R <sub>merge</sub> (%)*	13.7 (83.8)
I/ $\sigma$ I*	13.64 (2.58)
Completeness (%)*	99.2 (90.4)
Redundancy*	7.17 (6.77)

**Table 5:** Summary of refinement statistics.

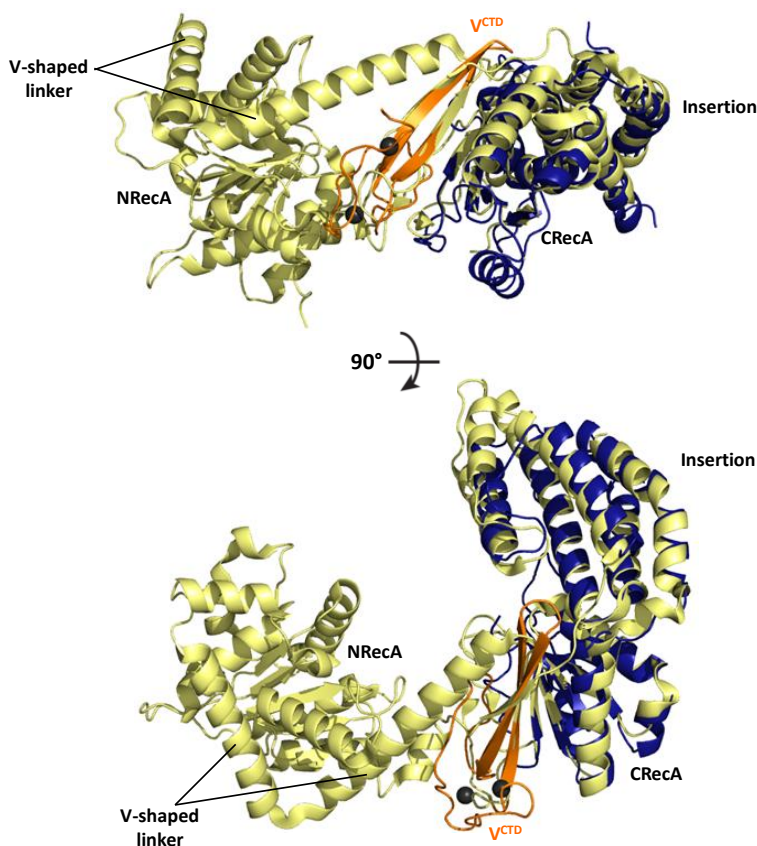
<b>Refinement</b>	
<b>Resolution</b>	48.14-2.29
<b>No. of Reflections</b>	27853
<b>R<sub>work</sub>/R<sub>free</sub> (%)</b>	17.92/22.96
<b>No. of atoms</b>	
<b>Protein</b>	2377
<b>Zn<sup>2+</sup></b>	2
<b>Water</b>	110
<b>B-factors</b>	
<b>Protein</b>	41.96
<b>Zn<sup>2+</sup></b>	51.12
<b>Water</b>	39.74
<b>R.m.s deviations</b>	
<b>Bond length (Å)</b>	0.009
<b>Bond angles (°)</b>	1.173
<b>Ramachandran (residues / %)</b>	
<b>favoured</b>	269 /95.7
<b>allowed</b>	12 / 4.3
<b>outlier</b>	0 / 0
<b>PDB code</b>	4I1S

Along with the identification of the viral zinc binding motif it becomes evident how the PIV5 V protein binds to its target MDA5. The typical C-terminal RecA-like fold of ATPases is characterized by a central  $\beta$ -sheet with parallel strands in the order  $\beta 3$ - $\beta 4$ - $\beta 2$ - $\beta 5$ - $\beta 1$ - $\beta 6$  (see figure 27) (24). In the determined structure, the CTD of the V protein forms a continuous  $\beta$ -sheet with the C-terminal RecA-like fold of MDA5 in the order  $\beta 3^{\text{MDA5}}$ - $\beta 4^{\text{MDA5}}$ - $\beta 2^{\text{MDA5}}$ - $\beta 5^{\text{MDA5}}$ - $\beta 1^{\text{VCTD}}$ - $\beta 2^{\text{VCTD}}$  (see figure 28). This arrangement implies that the V protein integrates itself into the structure of MDA5. Apparently, it melts the central  $\beta$ -sheet of the RecA-like fold and replaces its canonical  $\beta$ -strands 1 and 6 by the two  $\beta$ -strands of the zinc binding motif. In this way, the CTD of the V protein creates a mimic of the C-terminal RecA-like fold that hampered its identification in the structure (see figure 27).

Superposition of the MDA5:V protein core complex and the complete RIG-I SF2 ATPase domain (see figure 29) implies that the insertion of the V protein is for steric reasons incompatible with the canonical arrangement and tertiary structure of the MDA5 SF2 ATPase domain. In absence of the V protein the two conserved helices of the V-shaped linker mediate contacts between the N- and C-terminal RecA-like fold (see also section 1.3.1) (26, 29, 30, 60). The displacement of the MDA5  $\beta$ -strand 1 (which forms a connection between the N-terminal RecA like fold and the  $\alpha$ -helical insertion (30)) and the



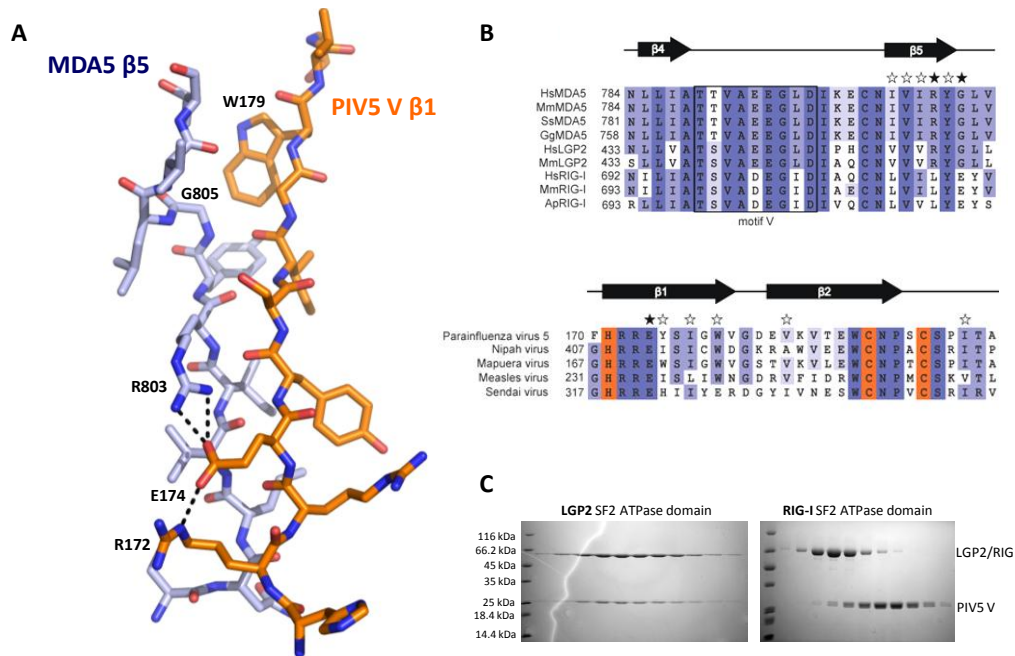
MDA5  $\beta$ -strand 6 (which lies directly upstream of the V-shaped linker (30)) by the CTD of the V protein is assumed to interfere with these interactions. The presence of the V protein is believed to severely disturb the normal architecture of the MDA5 SF2 ATPase domain and assumed to cause a disruption or partial unfolding of the ATPase domain. Such a scenario might explain why the MDA5:V protein complex did not crystallize (see sections 3.1-3.3) or solely crystallized in the presence of a protease that allowed cleavage of possible flexible portions. These portions include the N-terminal RecA-like fold, the  $\beta$ -strands 1 and 6 of the C-terminal RecA-like fold and the V-shaped linker of the ssMDA5 protein as well as the N-terminal domain (NTD) of the PIV5 V protein. All of these were present in the protein constructs used for crystallization but not in the determined crystal structure.



**Figure 29:** Superposition of the ssMDA5:PIV5 V core complex and the mmRIG-I SF2 ATPase domain (PDB code: 3TBK) with annotated domains and structural elements. The mmRIG-I SF2 ATPase domain is shown in yellow, ssMDA5 parts in dark blue and the PIV5 V C-terminal domain (CTD) in orange. Zinc ions are highlighted in grey. The N-terminal RecA-like fold, the  $\beta$ -strands 1 and 6 of the C-terminal RecA-like fold and the V-shaped linker of the ssMDA5 protein as well as the N-terminal domain (NTD) of the PIV5 V protein are not present in the structure, presumably because of proteolytic cleavage.

### 3.5 Differences in residue conservation enable RIG-I to evade V protein recognition

Figure 30A shows the interactions within the interface between  $\beta$ -strand 5 of the ssMDA5 C-terminal RecA-like fold and  $\beta$ -strand 1 of the PIV5 V protein CTD in detail. The interface is dominated by hydrophobic interactions, but E 174<sup>PIV5</sup> forms sequence-specific hydrogen bonds and ion pair interactions with R803<sup>ssMDA5</sup>. E174<sup>PIV5</sup> as well as R803<sup>ssMDA5</sup> are highly conserved among V proteins and MDA5 proteins, respectively (see figure 30B). Whereas LGP2 protein shares the conservation of R803<sup>ssMDA5</sup>, RIG-I proteins possess a conserved leucine residue at this position (see figure 30 B). These findings explain why paramyxoviral V proteins are able to bind MDA5 as well as LGP2 across species barriers, but not RIG-I (see figures 23 and 30C). The leucine residue allows RIG-I proteins to evade recognition by paramyxoviral V proteins as it provides no anchoring point for salt bridge formation with the viral glutamate residue.



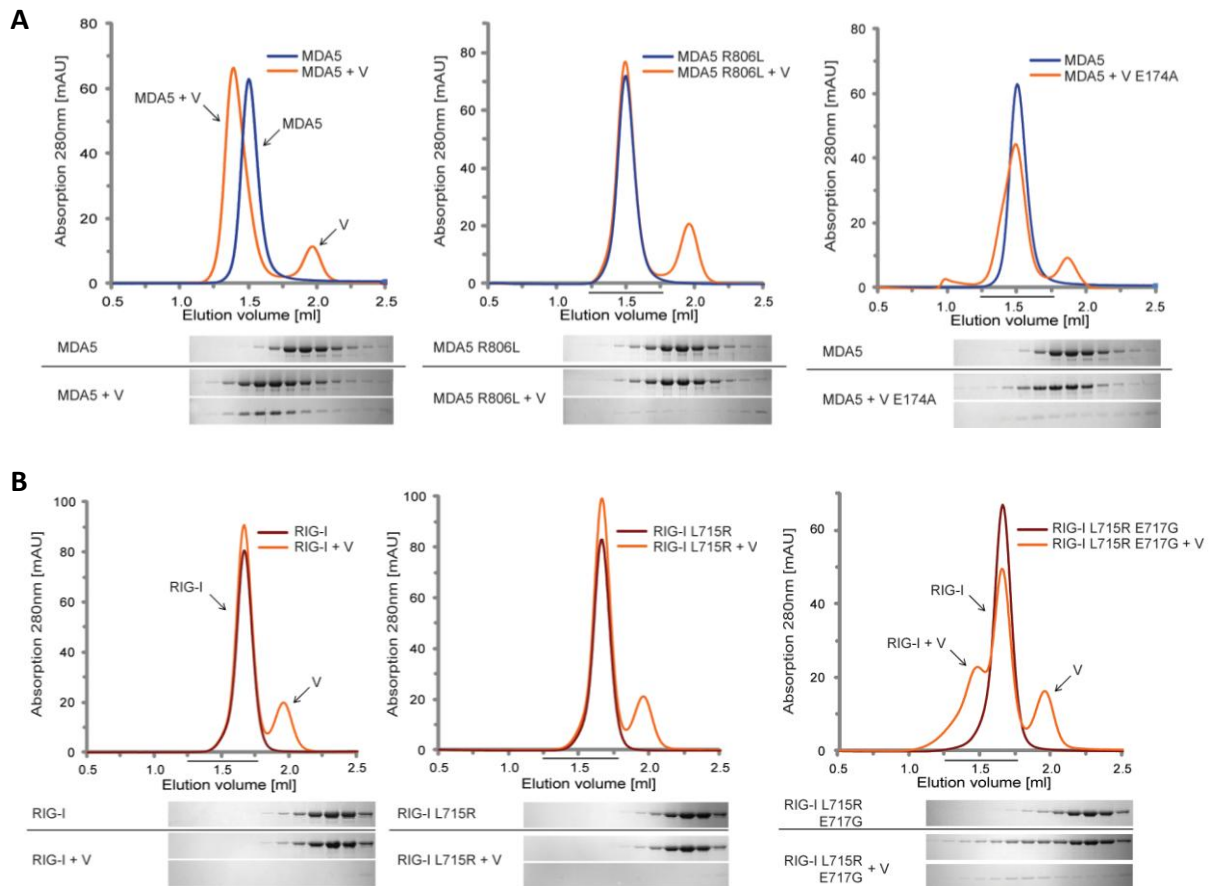
**Figure 30:** (A) Close-up views of the intermolecular ssMDA5:PIV5 V interactions. Salt bridge formation is indicated by dashed lines. SsMDA5 parts are shown in blue and PIV5 V protein parts in orange. (B) Structure-based alignments of the interface sequences of selected RIG-I like receptors (upper panel) and paramyxoviral V proteins (lower panel) with highlighted conserved motifs and residues (according to percentage identity). Stars indicate residues involved in interface interactions; filled stars indicate residues mutated in this study. (C) Stoichiometric coelution of the ssLGP2 SF2 ATPase domain (left panel) and separate elution of the mmRIG-I SF2 ATPase domain (right panel) and PIV5 V protein in size exclusion chromatography analysed by SDS-PAGE and Coomassie staining.



The ssMDA5:PIV5 V protein structure and sequence alignments revealed a second important difference in residue conservation among RLR family members that has critical impact on V protein binding. G805<sup>ssMDA5</sup>, that is highly conserved among MDA5 and LGP2, stacks against W179<sup>PIV5</sup> and allows it to insert into a hydrophobic pocket (see figure 30A). Instead of the glycine residue, RIG-I proteins have a conserved glutamate residue at this position, which sterically hinders the insertion of the conserved viral tryptophan residue (see figure 30B).

To confirm the interface observed in the ssMDA5:PIV5 V protein crystal structure and to approve importance of the found interactions, mutants of full length mmMDA5 as well as full length PIV5 V protein were generated and tested for their ability to form stable complexes in size exclusion chromatography (see figure 31A). Whereas wild type mmMDA5 shows stoichiometric coelution with wild type PIV5 V protein (see also section 3.1), a single amino acid exchange of R806<sup>mmMDA5</sup> (equivalent to R803<sup>ssMDA5</sup>) with leucine (as found in RIG-I) completely abolishes the interaction in size exclusion chromatography. Likewise, mutation of E174<sup>PIV5</sup> to alanine strongly reduces the *in vitro* binding. These studies verify that the structurally identified salt bridge is crucial for V protein binding to MDA5 and indicate evolutionary conservation of the binding mode.

To test if the structure-derived model of MDA5 recognition by paramyxoviral V proteins is sufficient to allow implementation of V protein binding in RIG-I proteins, full length mmRIG-I was mutated and tested for V protein interaction (see figure 31B). Wild type mmRIG-I shows no binding of PIV5 V protein in size exclusion chromatography. A single amino acid exchange of L715<sup>mmRIG-I</sup> with MDA5's arginine was found to be insufficient to induce V protein binding. The introduction of a second mutation replacing E716<sup>mmRIG-I</sup> by MDA5's glycine assumed to release the steric clash with W179<sup>PIV5</sup> led to interaction with PIV5 V protein that was partly stable in gel filtration. Thus, the engineered RIG-I double mutant validates the structural findings by gain of function.



**Figure 31:** Analytical size-exclusion chromatograms of mmMDA5, mmRIG-I, PIV5 V protein and mutants as indicated. Starting concentration is 15  $\mu$ M for MDA5/RIG-I and 22.5  $\mu$ M for PIV5 V protein. Elution profiles for MDA5 are in blue, for RIG-I in red and for MDA5- or RIG-I-V-mixtures in orange. Solid lines mark fractions analyzed by SDS-PAGE and Coomassie staining. The coelution of proteins indicates their interaction. **(A)** Mutations affecting the salt bridge between R806 of mmMDA5 and E174 of PIV5 V protein abolish or significantly impair their interaction. **(B)** Detectable binding of PIV5 V protein to mmRIG-I in size exclusion chromatography requires two mutations within the RIG-I SF2 ATPase domain: L715R is thought to enable ion pairing with E174 of PIV5 V protein and E717G is thought to reduce steric hindrance within the RIG-I:V protein interface.

### 3.6 The PIV5 V protein has to unfold itself to disrupt the fold of MDA5

In the crystal structure of PIV5 V protein bound to DDB1 (104) (the substrate recruiting adaptor protein of the ubiquitin ligase Cul4A, see also sections 1.5.1 and 4.4) the two  $\beta$ -strands of the C-terminal zinc binding motif form a continuous  $\beta$ -sheet with the NTD (see figure 32A). This central  $\beta$ -sheet is sandwiched by one  $\alpha$ -helix and long loops of the NTD. Provided that the isolated unbound form of the

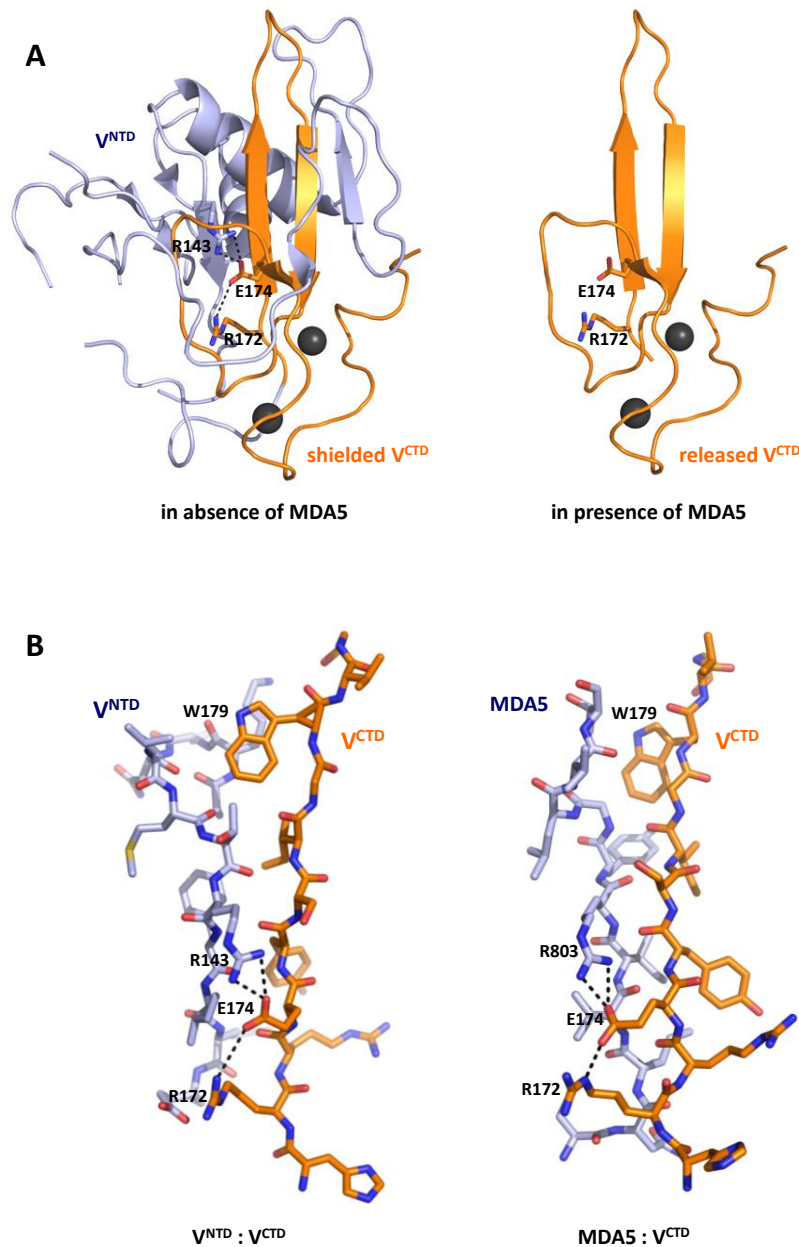
PIV5 V protein adopts a similar structure, the V protein must undergo a large conformational change to release the CTD for binding to MDA5. These observations suggest that the PIV5 NTD might shield the CTD in absence of MDA5, but exposes the CTD in presence of MDA5 (see figure 32). Whereas the two  $\beta$ -strands of the exposed CTD incorporate into the C-terminal RecA-like fold of MDA5 in order to disrupt its SF2 ATPase domain, the free NTD might adopt the unfolded conformation found in P protein (106).

Figure 32B compares the intermolecular ssMDA5:PIV5 V<sup>CTD</sup> and the intramolecular PIV5 V<sup>NTD</sup>:V<sup>CTD</sup> interactions observed for the PIV5 V protein in complex with DDB1 (104). Both interfaces share high similarity. The NTD of the PIV5 V protein even mimics the salt bridge observed in the MDA5:PIV5 interface via R143<sup>PIV5</sup> that binds to E174<sup>PIV5</sup> (see figure 32B). The NTD also provides a hydrophobic pocket for W179<sup>PIV5</sup>. This implies, however, that all these interactions between the NTD and CTD have to be relieved to allow MDA5 binding.

In agreement with the assumed unfolding processes of the MDA5 SF2 ATPase domain and the PIV5 V protein upon complex formation are results of small angle X-ray scattering (SAXS) experiments performed with mmMDA5 and mmMDA5:PIV5 V protein complexes (see figure 33). SAXS measures the distribution of intermolecular distances in proteins in solution. The pair distance distribution function (see also section 2.2.8) can be understood as a two-dimensional Patterson function that describes the paired set of all distances between all electrons within a protein. The position where the pair distance distribution returns to zero represents the maximum length in the particle ( $D_{\max}$ ). The pair distribution functions and *ab initio* models demonstrated a distinct increase in  $D_{\max}$  of the mmMDA5 full length molecule as well as the SF2 ATPase domain in presence of the PIV5 V protein (see figure 33A and B, see also figure 19 of section 3.2). Without knowledge of structural details, it was assumed that paramyxoviral V proteins recognize a conserved surface patch of MDA5 and bind on top of its SF2 ATPase domain resulting in an increased  $D_{\max}$ . However, the determined structure of the MDA5:V protein core complex revealed that V proteins rather recognize a conserved motif buried in the SF2 ATPase domain. The observed expansion of the MDA5 molecule and the increased  $D_{\max}$ -values apparently reflect wedging of the C-terminal zinc binding motif into the  $\beta$ -sheet of MDA5's C-terminal RecA-like fold and loosening of interactions within the SF2 ATPase domain (see section 3.4 and figure 33C).

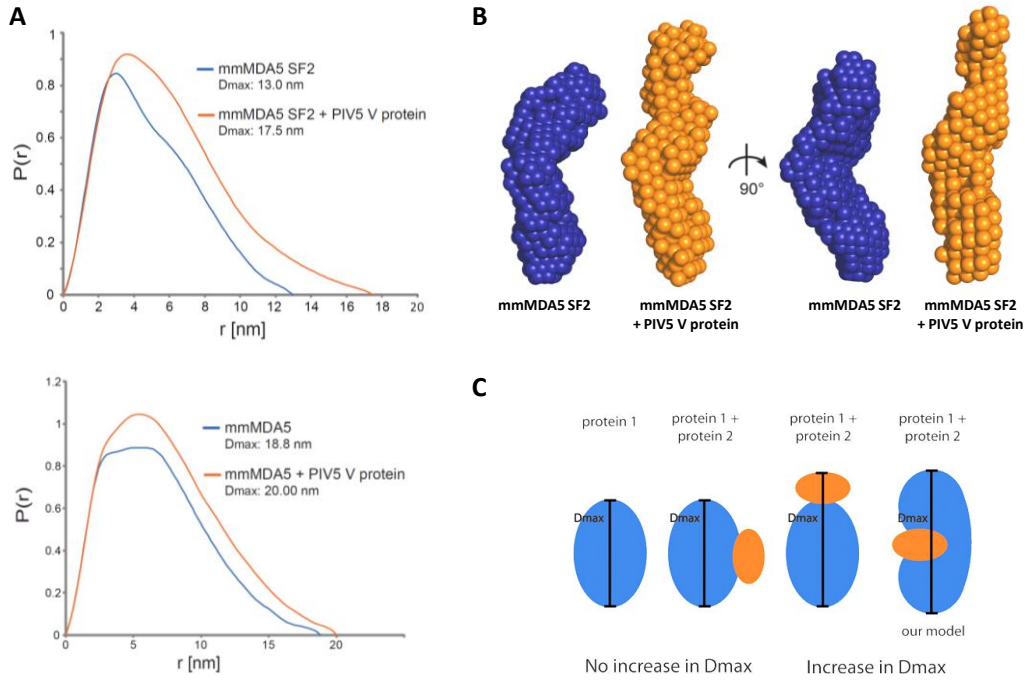
Differential scanning fluorimetry experiments examining the thermostability of purified proteins revealed a significant increase by approximately 10°C of the transition midpoint of different MDA5 protein constructs in presence of the PIV5 V protein (see figure 34). These data imply that the conserved core of the MDA5:V protein complex is highly stable. For the PIV5 V protein itself it was not possible to determine a reasonable transition midpoint of protein unfolding indicating structural instability of the PIV5 V protein in isolation. Formation of a highly stable core within the MDA5:V protein complex might

favor V protein binding to MDA5 and this stabilization effect might outbalance negative effects of potentially unfolding processes (see section 3.4).

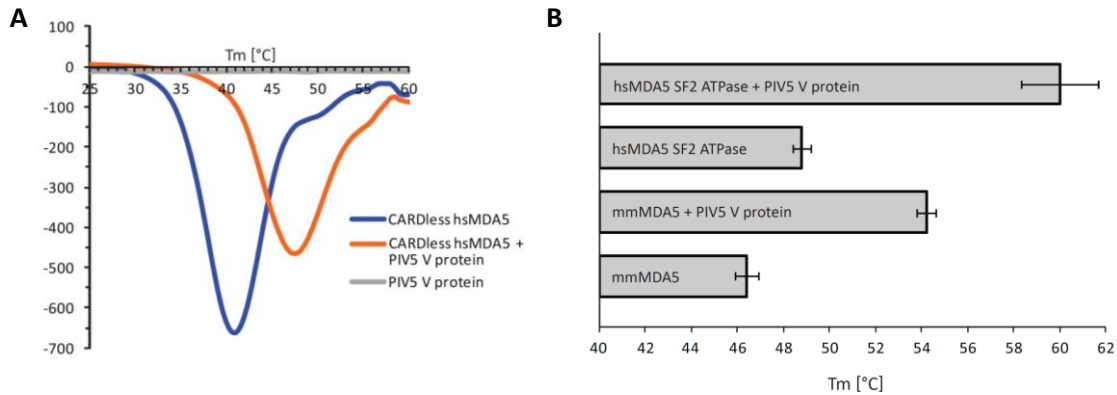


**Figure 32:** (A) Crystal structure of the PIV5 V protein in complex with DDB1 (not shown here; PDB code: 2HYE). The N-terminal domain (NTD, shown in light blue) has to release the C-terminal zinc binding motif (CTD, shown in orange) to allow binding to MDA5. (B) Close-up views of the intramolecular PIV5  $V^{NTD} : V^{CTD}$  (left panel) and the intermolecular ssMDA5:PIV5  $V^{CTD}$  interactions (right panel). Salt bridges are indicated by dashed lines. Parts of ssMDA5 and of the PIV5 NTD are shown in light blue. Parts of the PIV5 CTD are shown in orange.

### III RESULTS



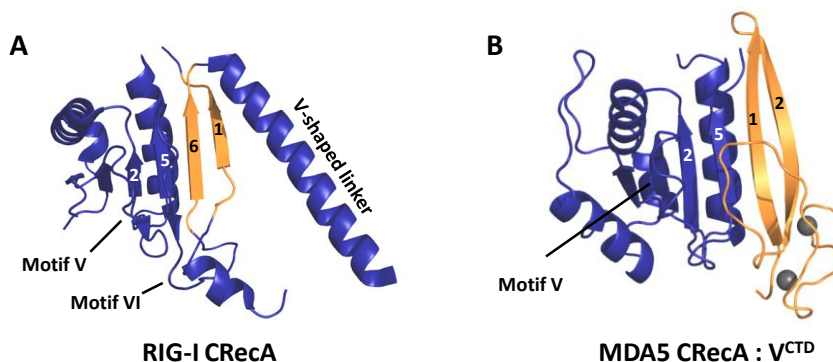
**Figure 33:** (A) Pair-distribution functions obtained by GNOM (138) of full length mmMDA5 (lower panel) or mmMDA5 SF2 ATPase domain (upper panel) in absence (shown in blue) and presence of PIV5 V protein (shown in orange) with indicated maximum particle sizes ( $D_{max}$ ). (B) Final averaged *ab initio* models of the mmMDA5 SF2 ATPase domain in isolation (shown in blue) and in complex with PIV5 V protein (shown in orange). (C) Possible scenarios that cause an increase in  $D_{max}$  in SAXS. The last panel represents the assumed model.



**Figure 34:** Thermostability of MDA5 in isolation and in complex with PIV5 V protein. Binding of the V protein increases the thermostability of different MDA5 constructs and species (mouse (mm) MDA5 full length, human (hs) MDA5 CARDless, human (hs) MDA5 SF2 ATPase domain) by approximately 10°C. The differential scanning fluorimetry experiments are either illustrated as melting curves (A) or as bars (B). The bars represent mean melting temperatures and the error bars indicate the standard deviation from five independent measurements.

### 3.7 V protein binding has dual impact on the function of MDA5

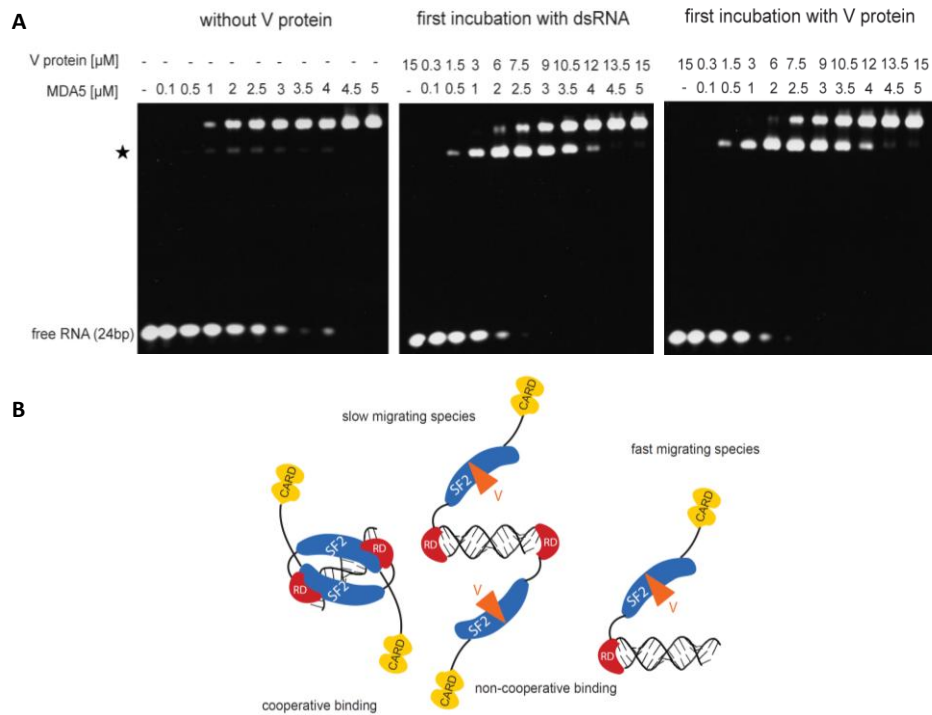
The insertion of the paramyxoviral C-terminal zinc binding motif displaces not only  $\beta$ -strand 1 and  $\beta$ -strand 6 of MDA5's C-terminal RecA-like fold, but also motif VI. This conserved motif, which MDA5 shares with all other members of the SF2 helicase family and is known to be essential for ATP binding and hydrolysis (24), is normally located between  $\beta$ -strand 5 and  $\beta$ -strand 6 (see figure 35) at the interface between the N- and C-terminal RecA-like folds. The crystal structure of the RIG-I SF2 ATPase domain in complex with an ATP transition state analog suggests that arginine residues of motif VI could bind the  $\gamma$ -phosphate of ATP (60). It was further proposed that one of the arginine residues within motif VI functions as an "arginine finger" that could stabilize the water-Mg<sup>2+</sup>- $\beta/\gamma$  phosphate intermediate in the course of the ATP hydrolysis reaction (24). The displacement of motif VI and general interference with the normal and functional architecture of the SF2 ATPase domain (see section 3.4) explain why paramyxoviral V proteins are able to abrogate the ATP hydrolysis activity of MDA5 (see section 1.5.3 and figure 8).



**Figure 35:** (A) Comparison of the functional C-terminal RecA-like fold of mmRIG-I (PDB code: 3TBK) and (B) the C-terminal RecA-like fold of ssMDA5 hijacked by the C-terminal domain (CTD) of PIV5 V protein with annotated  $\beta$ -strands and conserved motifs.

To investigate if V protein binding also interferes with the RNA binding activity of MDA5, electrophoretic mobility shift assays (EMSAs) with full length mmMDA5 were performed in absence and presence of PIV5 V protein (see figure 36A). These experiments revealed that the PIV5 V protein induces substantial changes of the MDA5 RNA-binding mode. In absence of V protein, mmMDA5 shifted a 24bp RNA predominantly to a single species with a slight amount of a faster migrating species at lower MDA5 concentrations. Presence of the V protein caused a strong accumulation of the faster migrating species irrespective of whether the first incubation step of MDA5 was performed with RNA or V protein. One

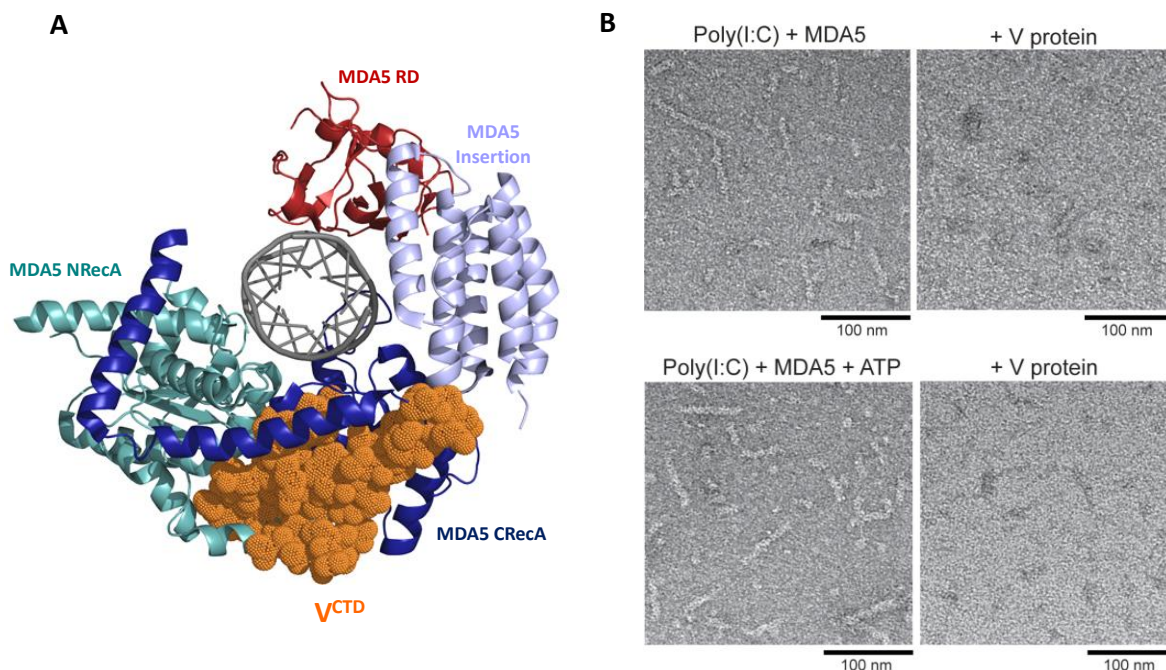
plausible explanation for the observed alteration in migration is that V protein binding interferes with the cooperative RNA-binding mode of MDA5. The three-dimensional reconstruction of RNA-bound MDA5 filaments by electron microscopy showed that MDA5 uses intermolecular contacts which involve all three subdomains of the SF2 ATPase domain to assemble along dsRNA (61) (see also section 1.3.2 and figure 6). In absence of the V protein, suchlike interactions could allow cooperative binding of two MDA5 molecules to the 24bp RNA-ligand resulting in one major, slowly migrating species. As V protein binding is assumed to disrupt the fold of the MDA5 SF2 ATPase domain, cooperative binding via intermolecular contacts of the ATPase subdomains might be hindered in presence of the V protein. The isolated MDA5 RD has been shown to bind RNA independently from the SF2 ATPase domain (56, 59). According to the crystal structure of the ssMDA5:PIV5 V protein core complex, this domain seems not to be affected by V protein binding. In presence of the V protein, an intact RD could allow MDA5 molecules to bind non-cooperatively to the ends of the dsRNA-ligand resulting in a mixture of fast migrating 1:1 and slowly migrating 2:1 MDA5:dsRNA complexes (see figure 33B).



**Figure 36:** (A) Electrophoretic mobility shift assays (EMSAs) of mmMDA5 incubated with palindromic 24bp RNA (1 μM final) in absence and presence of PIV5 V protein. Incubation of mmMDA5 with dsRNA after or before addition of PIV5 V protein does not change the observed pattern in migration. The fast migrating species is indicated by a star. (B) Models of possible MDA5:dsRNA species observed in EMSAs in absence and presence of paramyxoviral V protein. Model for cooperative binding as described in (54, 61).



Superposition of the C-terminal zinc binding motif and the MDA5:RNA clamp structure (see figure 37A) illustrates that wedging of the CTD into the SF2 ATPase domain disrupts structural patterns that are essential for binding of RNA and required for the formation of a proper MDA5:MDA5 interface within RNA-bound filaments (61) (see also section 1.3.2). To directly prove that V protein binding indeed interferes with the cooperative RNA-binding mode of MDA5, filament formation of full length MDA5 was analyzed in absence and presence of V protein employing electron microscopy. Figure 37B shows negatively stained filaments formed by purified mmMDA5 on poly(I:C). The presence of PIV5 V protein drastically reduced the number of filaments independent from the absence or presence of ATP. Thus, the paramyxoviral V protein is able to affect both, the ATPase activity as well as the cooperative RNA-binding activity of MDA5.



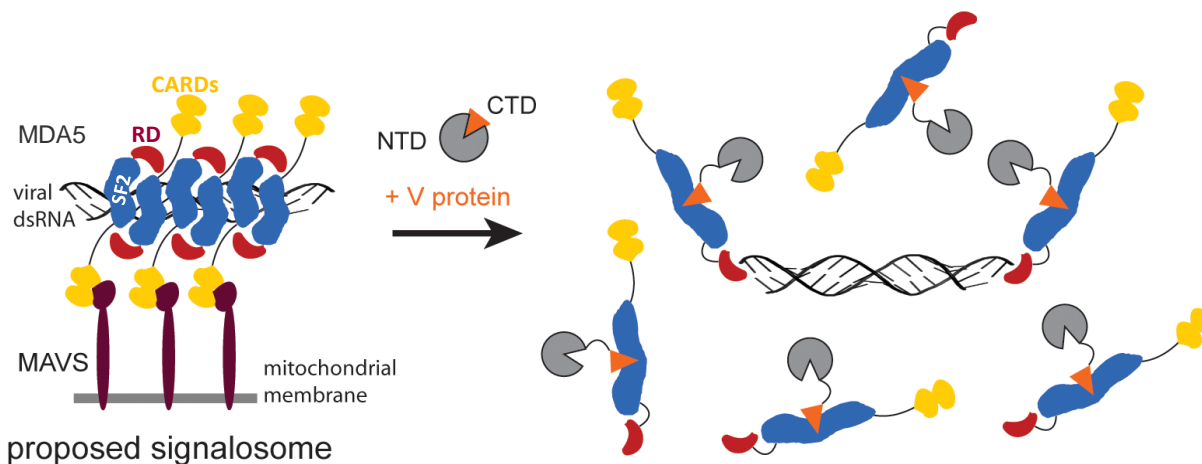
**Figure 37:** (A) Structure of human CARDless MDA5 in complex with dsRNA (PDB code: 4GL2) with superimposed PIV5 V<sup>CTD</sup> (dot density shown in orange and modeled from the position in the ssMDA5:PIV5 V protein complex). The subdomains of the MDA5 SF2 ATPase domain are shaded in blue, the RD is shown in red and the dsRNA-ligand in grey. (B) Filament formation of mmMDA5 on poly(I:C) visualized by negative-stain electron microscopy in the presence and absence of ATP and PIV5 V protein. Shown are representative grid areas.



## 4. Discussion

### 4.1 A structure-derived model for MDA5 inhibition by paramyxoviral V proteins

The integration of the findings presented in this thesis with the current state of knowledge in the field of RLR function and signaling allows us to derive of a first molecular mechanism for the inhibition of the innate immune receptor MDA5 by paramyxoviral V proteins (see figure 38). In the absence of V protein, MDA5 is able to hydrolyze ATP and to cooperatively assemble into filaments along viral dsRNA (13, 30, 53, 54, 61) (see also section 1.2.1). This ensures the formation of an intact and functional signalosome that involves homotypic CARD-CARD interaction with the signaling adaptor protein MAVS and its polymerization into fibrils on the mitochondrial membrane (30, 67-73) (see also section 1.4.1). The MDA5-MAVS signaling entity is believed to be crucial for the activation of pathways that finally induce an antiviral state of the host cell (4, 5, 66, 76, 77, 79) (see also section 1.4.2).

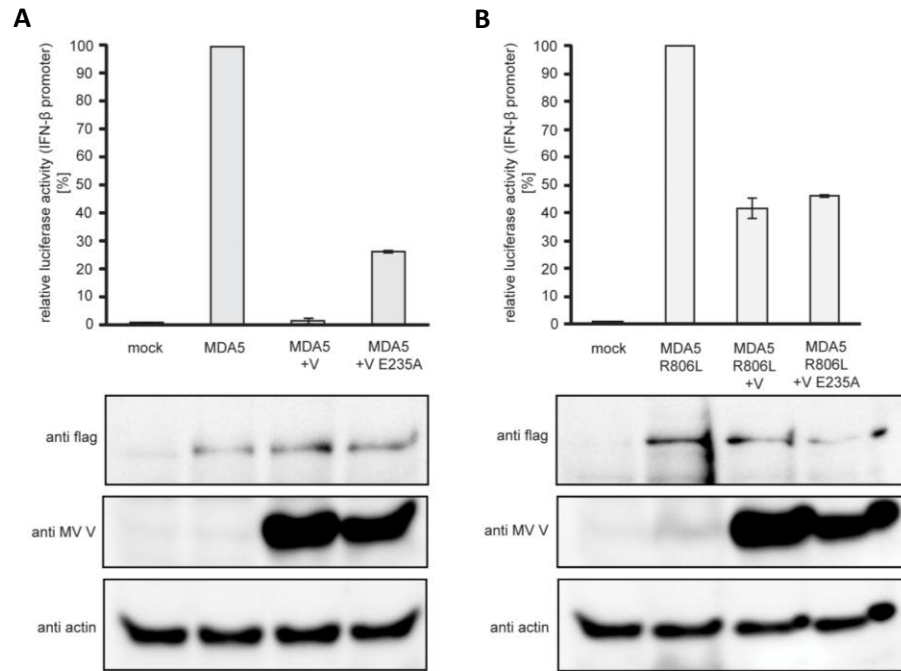


**Figure 38:** Proposed model for inhibition of MDA5 signaling assemblies by paramyxoviral V proteins. The dsRNA-bound MDA5 filaments are assumed to interact with MAVS and induce the formation of MAVS fibrils postulated to be active signal-entities (30, 67-73). The V protein disrupts the architecture of the MDA5 SF2 ATPase domain by a  $\beta$ -strand replacement mechanism and is assumed to unfold itself in the process of MDA5 binding. V protein-induced partial unfolding of MDA5 results in disruption of functional MDA5:dsRNA filaments and hence signal-transmission of the antiviral response. The structure and interactions of the NTD in the MDA5 complex are only illustrated and need to be addressed in future studies. Model for cooperative RNA binding by dsMDA5 as described in (54, 61).

The structural determination of the conserved MDA5:V protein complex (131, 134-136) revealed that paramyxoviral V proteins are able to trap MDA5 in an inactive state by disrupting its SF2 ATPase domain. The V protein inserts its C-terminal zinc binding motif into the central  $\beta$ -sheet of the MDA5 C-terminal RecA-like fold and displaces its canonical  $\beta$ -strands 1 and 6 (see section 3.4). This partial unfolding of the innate immune receptor abolishes its ATPase activity and simultaneously disturbs cooperative filament formation on dsRNA presumably by breaking up the MDA5:MDA5 interface (54, 61) as well as the MDA5:dsRNA interface (30) (see section 3.7). It has been shown that ATP hydrolysis and filament formation by MDA5 have to go hand in hand to assemble a functional MDA5-MAVS signalosome and to trigger downstream signaling cascades that limit viral replication (5, 30, 66, 78, 83-86) (see also sections 1.3 and 1.4). Targeting both activities, the paramyxoviral V protein represents a very potent inhibitor of MDA5 that can antagonize MDA5-mediated host immune response at the very first step of viral ligand recognition. Unlike in the case of paramyxoviral inhibition of common downstream TLR/RLR signaling components (e.g. IKK $\epsilon$ /TBK1, NF- $\kappa$ B, IRF3 and IRF7, STAT1 and STAT2; see also figure 6 and section 1.5.2), this enables the V protein to avoid amplification of the signal during the transduction cascade and hence efficient repression of MDA5 pathways.

Experiments examining the activation of a interferon- $\beta$  promoter responsive to the transcription factors AP-1, NF- $\kappa$ B, IRF-3 and IRF-7 (153) by human MDA5 in living cells emphasize the importance of MDA5 as target for paramyxoviral V proteins (see figure 39). In these gene reporter assays, the signaling activity of wild type hsMDA5 was found to be almost completely inhibited in the presence of wild type MV V protein, whereas the E235A<sup>MV</sup> mutant caused a distinct attenuation of the signaling inhibition (see figure 39A and also (131)). This mutant was observed to be impaired in its interaction with MDA5 in living cells (see figure 43 and also (131)) as well as in analytical size exclusion chromatography experiments (see figure 31A with the equivalent E174A<sup>PIV5</sup> mutant), indicating that the partial reconstitution of the MDA5 signaling pathway seen in the interferon- $\beta$  promoter dependent gene reporter assays is caused by restricted binding of the V protein mutant to MDA5. Similar reduction of V protein-mediated inhibition of interferon- $\beta$  promoter activity could be obtained with the structure-derived R806L<sup>hsMDA5</sup> mutant (see figure 39B), which was shown to be able to evade V protein recognition in living cells (see figure 43) as well as *in vitro* (see figure 31A with the equivalent R806L<sup>mmMDA5</sup> mutant). The measured degree of the partial rescue obtained with the R806L<sup>hsMDA5</sup> mutant was not found to be enhanced in the presence of the E235A<sup>MV</sup> mutant. This positive control verifies that the observed reconstitution of the MDA5 signaling pathway indeed results from impaired MV V protein binding to R806L<sup>hsMDA5</sup>. The incomplete rescue of the MDA5 signaling activity in both experiments is most likely the consequence of the multifunctionality of the V protein, which is known to inhibit the induction of

interferon- $\beta$  at several levels (e.g. at the level of the IRF3/7 activating kinase IKK $\epsilon$ /TBK1 and at the level of the transcription factors IRF3/7 as well as NF- $\kappa$ B) (109, 111-114) (see also section 1.5.2 and figure 6).



**Figure 39:** Attenuation of V protein mediated MDA5 signaling inhibition by mutation of the conserved salt bridge within the MDA5:V protein interface (57) (data of Dr. Kerstin Monika Schuhmann and Prof. Karl-Klaus Conzelmann, Gene Center Munich). Shown are interferon- $\beta$  promoter dependent luciferase reporter gene assays performed in HEK-293T cells with a promoter sequence, which is responsive to the transcription factors AP-1, NF- $\kappa$ B, IRF-3 and IRF-7 (153). Mean values ( $\pm$  SEM) were obtained from experiments performed in triplicate (N=3). Expression levels of transfected flag-tagged hsMDA5 and MV V protein plasmids were analyzed by western blot (lower panels). (A) Reporter gene assay performed with hsMDA5 WT and MV V WT or MV V E235A. Shown are relative values of the interferon- $\beta$  promoter activity related to the MDA5 WT value (100%). (B) Reporter gene assay performed with hsMDA5 R806L and MV V WT or MV V E235A. Shown are relative values of the interferon- $\beta$  promoter activity related to the MDA5 R806L value (100%).

#### 4.1.1 Which energy source drives the unfolding processes that accompany V protein binding?

Provided that the PIV5 V protein structure observed within the DDB1 complex (*104*) (see also sections 1.5.1 and 4.4) is stable in absence of a binding partner as well, the V protein has to unfold itself to allow disruption of the MDA5 SF2 ATPase fold (see section 3.6). According to this structure, the insertion of the viral C-terminal zinc binding motif (CTD) into the C-terminal RecA-like fold of MDA5 and formation of the MDA5:V protein interface requires not only abrogation of intermolecular interactions within the MDA5 molecule but also within the PIV5 V protein (see figure 32). This raises the question which energy source actually drives the unfolding processes postulated in this thesis. Differential scanning fluorimetry experiments suggest that a significant gain in thermostability might favor complexation of MDA5 and the PIV5 V protein (see figure 34). The experiments revealed that the melting temperature of isolated MDA5 in presence of the V protein ligand increases by approximately 10°C indicating that the formation of a highly stable MDA5:V protein core could compensate and outbalance the presumed unfolding processes. The thermal shift assays further revealed that the PIV5 V protein itself is not very stable as it did not show a reasonable transition midpoint of protein unfolding (see figure 34A). This finding argues for an undefined and transient tertiary structure of the PIV5 V protein and makes doubt that the structure observed within the DDB1 complex represents a stable conformation. This observation leads to the question whether V protein binding to MDA5 involves an induced fit mechanism of the viral N-terminal domain (NTD). The NTD of the PIV5 V protein is not present in the determined crystal structure of the ssMDA5:PIV5 V protein core complex, probably because it adopts the unfolded conformation seen in P protein (*106*) and its flexibility caused proteolytic cleavage during the crystallization process (see sections 3.3 and 3.4). However, the possibility that the NTD of the PIV5 V protein binds to MDA5 portions, which are missing in the crystal structure (namely CARDs, N-terminal RecA-like fold and RD), and thereby undergoes structural stabilization via an induced fit mechanism cannot be ruled out. As the PIV5 V protein has been shown to be able to bind the isolated N-terminal RecA-like fold of MDA5 in yeast two-hybrid assays (*136*) (see also section 1.5.2), this might represent a reasonable scenario. Analysis by NMR might be suitable to gain further insight into the nature of potential MDA5:V protein interactions that defied crystallization and possible associated induced fit or unfolding processes. Hydrogen-deuterium exchange might allow to compare the solvent accessibility of different protein parts in isolation and within the complex and hence to complement the MDA5:V protein binding mode presented in this thesis.

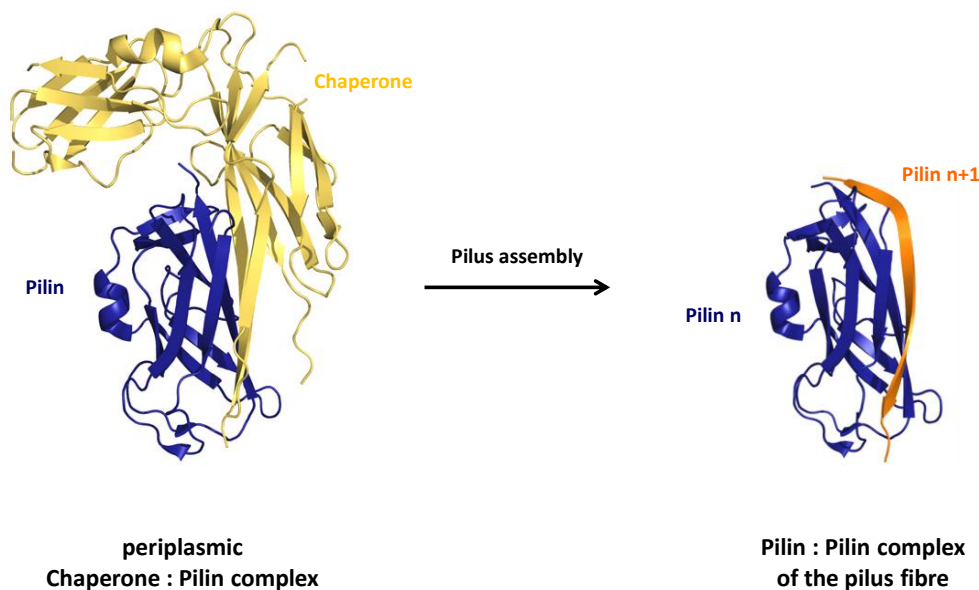
Moreover, it has to be considered that the observed and described gain in MDA5's thermostability refers to monomeric MDA5 proteins in an RNA-unbound state, but not to cooperatively assembled, RNA containing MDA5 filaments (see figure 34). In this context it would be interesting to

know, how the inhibition of MDA5 by paramyxoviral V proteins actually proceeds *in vivo*. It has not yet been fully elucidated if the V protein already exists within the virion and enters the host cell with the viral ribonucleoprotein (RNP) complex or if it is produced later during viral replication within the host cell (98, 102). The first scenario would allow instant prevention of MDA5 filament formation on viral dsRNA ligands by paramyxoviral V proteins and therefore efficient inhibition of MDA5 mediated antiviral signaling. The second scenario raises the question, if paramyxoviral V proteins possess the capability to dissolve already formed and robust MDA5 filaments. The EMSA experiments presented in this thesis (see section 3.7 and figure 36) indicate that V proteins might be able to remove MDA5 proteins from dsRNA. However, if the actual *in vivo* situation, in which MDA5 filaments might be stabilized or protected by other protein factors or specific subcellular localization (e.g. formation of cytoplasmic foci or docking to MAVS at the outer mitochondrial membrane (30, 67-73) (see also section 1.4.1)), allows such a process requires further investigation and better understanding of MDA5 filament assembly and signaling *in vivo*.

#### 4.1.2 Known $\beta$ -strand displacement mechanisms in other biological processes

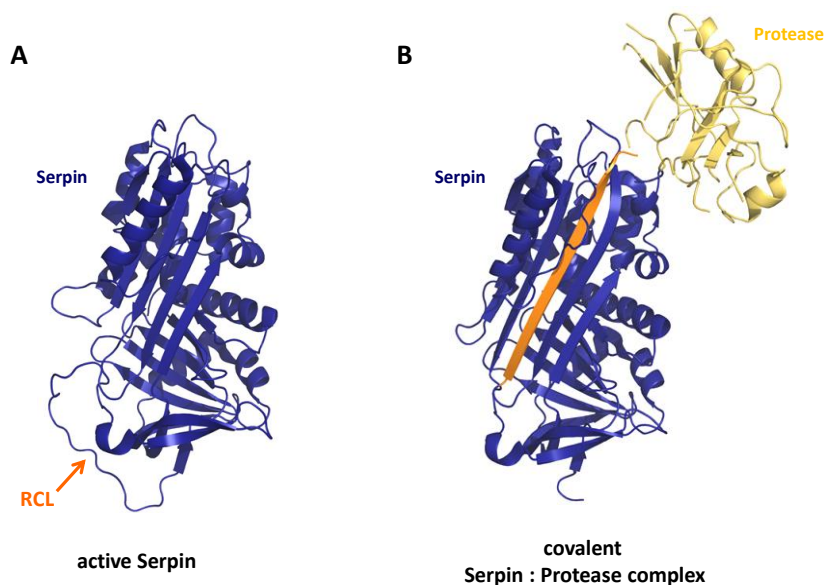
The regard of other known  $\beta$ -strand displacement mechanisms might contribute to improved comprehension of the complexation process between MDA5 and the V protein. A well studied  $\beta$ -strand displacement mechanism is the chaperone-assisted assembly of adhesive pili on the surface of gram-negative bacteria (see figure 40). These bacteria use their general secretory pathway to export pilin subunits (pilins) to the periplasm, where they are captured and stabilized by SafB chaperones. Pilins adopt an immunoglobulin-like (Ig-like) fold that lacks the seventh, C-terminal  $\beta$ -strand and thus exhibit an exposed hydrophobic groove on its surface. The SafB chaperone assists pilins in folding by providing the missing  $\beta$ -strand in *trans*. The donor-strand of the chaperone shields the hydrophobic groove, complements the Ig-like fold and simultaneously prevents premature polymerization of pilin subunits by occupying their interaction surface. Pilin assembly requires the formation of a transient ternary complex that involves the chaperone:pilin complex and the N-terminal  $\beta$ -strand of an incoming pilin subunit. In the course of the polymerization reaction, the donor-strand of the chaperone is replaced by the N-terminal  $\beta$ -strand of the next pilin subunit (154). This donor-strand exchange process was shown to proceed spontaneously *in vitro* (155) and to occur *in vivo* at the outer membrane of gram-negative bacteria independent from the presence of an energy source (156). The chaperone primes the pilin protein for a topological transition that drives pilin polymerization into fibers. The bound chaperone keeps the hydrophobic groove of the pilin subunit in an open, activated conformation. During the exchange of the

donor strand, the pilin subunit closes the groove and seals the N-terminal  $\beta$ -strand of the next subunit in place (155).



**Figure 40:** Structural comparison of the pilin subunit (shown in blue) in complex with the SafB chaperone (shown in yellow) (left panel, PDB code: 2CO6) and in complex with the N-terminal  $\beta$ -strand of an adjacent pilin subunit (shown in orange) of the pilus fibre (right panel, PDB code: 2CO4). Adapted from (154).

Structural work on serpins (serine protease inhibitors) revealed tremendous and irreversible conformational changes during their mechanism of action that involves a  $\beta$ -strand insertion. Serpins represent covalent “suicide” protease inhibitors and therefore crucial regulators of many and diverse biological processes. Serpins expose a reactive center loop (RCL) that provides a “bait” peptide bond for serine proteases (see figure 41). The nucleophilic attack by the serine residue of the catalytic triad located in the active site of the protease leads to the formation of a covalent complex between inhibitor and protease. The cleaved and now unconstrained RCL is inserted into a central  $\beta$ -sheet of the serpin protein and thereby traps the covalent complex. Similar to V protein-bound MDA5, the RCL-cleaved form of serpin inhibitors shows a significant higher thermostability as the native, uncleaved state indicating a global stabilization of the serpin protein within the covalent complex and explains why an active inhibitor cannot be recovered without providing energy (157).

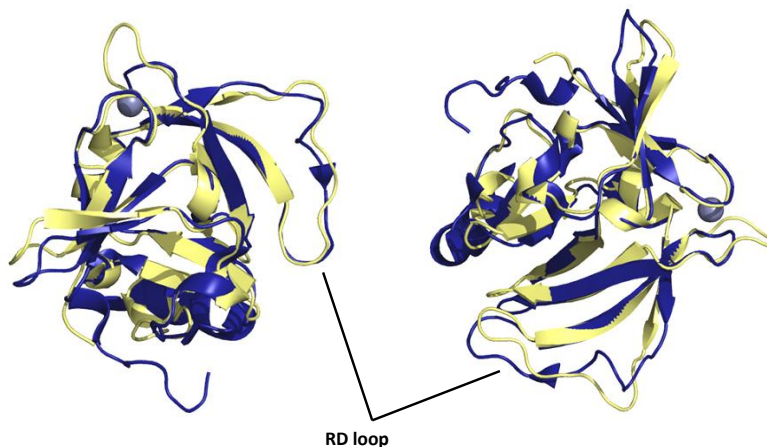


**Figure 41:** (A) Crystal structure of an active serpin (shown in blue). The inhibitory reactive center loop (RCL) is indicated by an arrow (PDB code: 1QLP). (B) Crystal structure of a covalent serpin:protease complex. The protease is shown in yellow, serpin in blue and the cleaved and inserted RCL in orange (PDB code: 1EZX). Adapted from (157-159).

#### 4.2 Is the RD of RLRs able to switch between two binding modes?

The V protein-induced disappearance of any filamentous structures of MDA5 on long dsRNA in EM experiments on the one hand, but clear binding of MDA5 to a short 24bp RNA ligand in EMSA experiments in the presence of the SF2 disruptive V protein on the other hand (see figures 37 and 36), raise the question if the MDA5 RD is able to switch between dsRNA stem recognition and dsRNA end recognition. The crystal structure of the MDA5 RNA clamp revealed that the MDA5 SF2 ATPase domain fixes dsRNA stem recognition by the MDA5 RD directly and indirectly (30) (see also section 1.3.2). In contrast, a conserved loop within the RIG-I RD mediates its binding to dsRNA ends (26, 29, 46) (see also sections 1.2.2 and 1.3.1). The equivalent loop of the MDA5 RD within the RNA clamp structure was found to be disordered due to displacement by the stem of the dsRNA ligand (30). However, the loop region of the free and isolated MDA5 RD is ordered and adopts a similar conformation as seen for the loop region of the RIG-I RD in the dsRNA-bound state (see figure 42) (30, 56). This observation implies that the MDA5 RD might have the potential to switch between two different dsRNA-binding modes. Moreover, the MDA5 RD loop is well conserved in sequence among MDA5 orthologs suggesting as yet undefined function of this region.

It has further been shown that RIG-I is able to translocate along dsRNA upon recognition of dsRNA ends (62), supporting the idea of plasticity of RD mediated dsRNA-binding. It would be interesting to address the question of different RD binding modes in further experiments and to elucidate possible functional impacts on RLR-mediated ligand sensing and antiviral signaling efficiency.



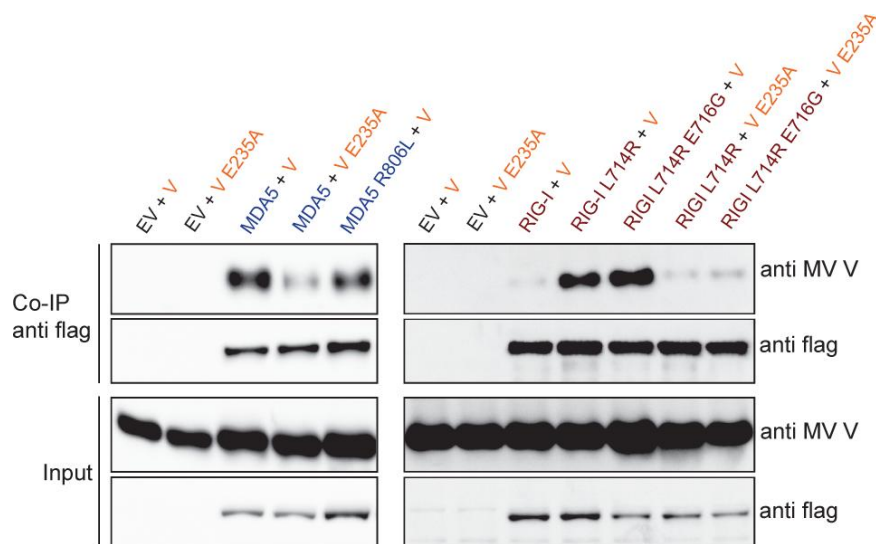
**Figure 42:** Superposition of the RNA-free MDA5 RD structure (shown in blue, PDB code: 3GA3) and the RNA-bound RIG-I RD structure (shown in yellow, PDB code: 3TMI) reveals similar conformation of the RD loop. Adapted from (29, 30, 56).

#### 4.3 V protein binding to other RLR family members and its functional impact

The crystal structure of the ssMDA5:PIV5 V protein core complex and *in vitro* validation experiments of the interface revealed a crucial role of salt bridge formation between a conserved arginine residue of MDA5 (that is missing in case of RIG-I) and a conserved glutamate residue of paramyxoviral V proteins for complexation (see section 3.5 and figures 30 and 31A). Consistent with these findings are results of co-immunoprecipitation experiments using flag-tagged hsMDA5 and MV V protein overexpressed in living cells (see figure 43, left panel) (57). Mutating R806<sup>hsMDA5</sup> (equivalent to R803<sup>ssMDA5</sup>) to a leucine residue (as found in RIG-I) reduced the interaction between hsMDA5 and MV V protein. Likewise, mutating E235<sup>MV</sup> (equivalent to E174<sup>PIV5</sup>) strongly reduced binding to hsMDA5 in living cells. These data verify the demonstrated importance of both sides of the salt bridge for MDA5:V protein complexation (see figure 31A). The *in vitro* experiments presented in section 3.5 and these co-immunoprecipitation experiments validate the MDA5:V protein interaction observed in the crystal structure (see figure 30A)



with proteins of four different species (mmMDA5, hsMDA5, PIV5 V protein, MV V protein) and therefore confirm evolutionary conservation of the binding mode.



**Figure 43:** Western-blot analysis of co-immunoprecipitation experiments performed with flag-tagged hsMDA5 or hsRIG-I and MV V protein expressed in HEK-293T cells (57) (data of Dr. Kerstin Monika Schuhmann and Prof. Karl-Klaus Conzelmann, Gene Center Munich). EV = empty vector. Left panel: Mutations affecting the salt bridge formation between the conserved glutamate of V protein and arginine of MDA5. Right panel: Mutation(s) of RIG-I to mimic the critical MDA5 arginine.

Moreover, the induced *in vitro* binding of PIV5 V protein to the structure-derived L715R E717G<sup>mmRIG-I</sup> mutant that mimics the critical MDA5 arginine residue (see section 3.5 and figure 31B) was validated in living cells by co-immunoprecipitation experiments using flag-tagged hsRIG-I and MV V protein (see figure 43, right panel) (57). In these experiments, the single exchange of L714<sup>hsRIG-I</sup> with MDA5's arginine was found to be sufficient to induce binding of MV V protein to hsRIG-I. The double mutant L714R E716G<sup>hsRIG-I</sup> (the second mutation accounts for the likely clash between a conserved tryptophan residue of the V protein and the glutamate that RIG-I has instead of the glycine found in MDA5, see also section 3.5 and figure 30) strengthens the observed interaction. This induced interaction was found to be in turn abolished by the E235A<sup>MV</sup> mutant. These data confirm that a different residue conservation within the C-terminal RecA-like fold of RIG-I enables this RLR family member to evade binding by paramyxoviral V proteins (134) (see also section 1.5.3) and that structure-guided exchange of only two amino acids is able to annul this evasion (see also section 3.5).

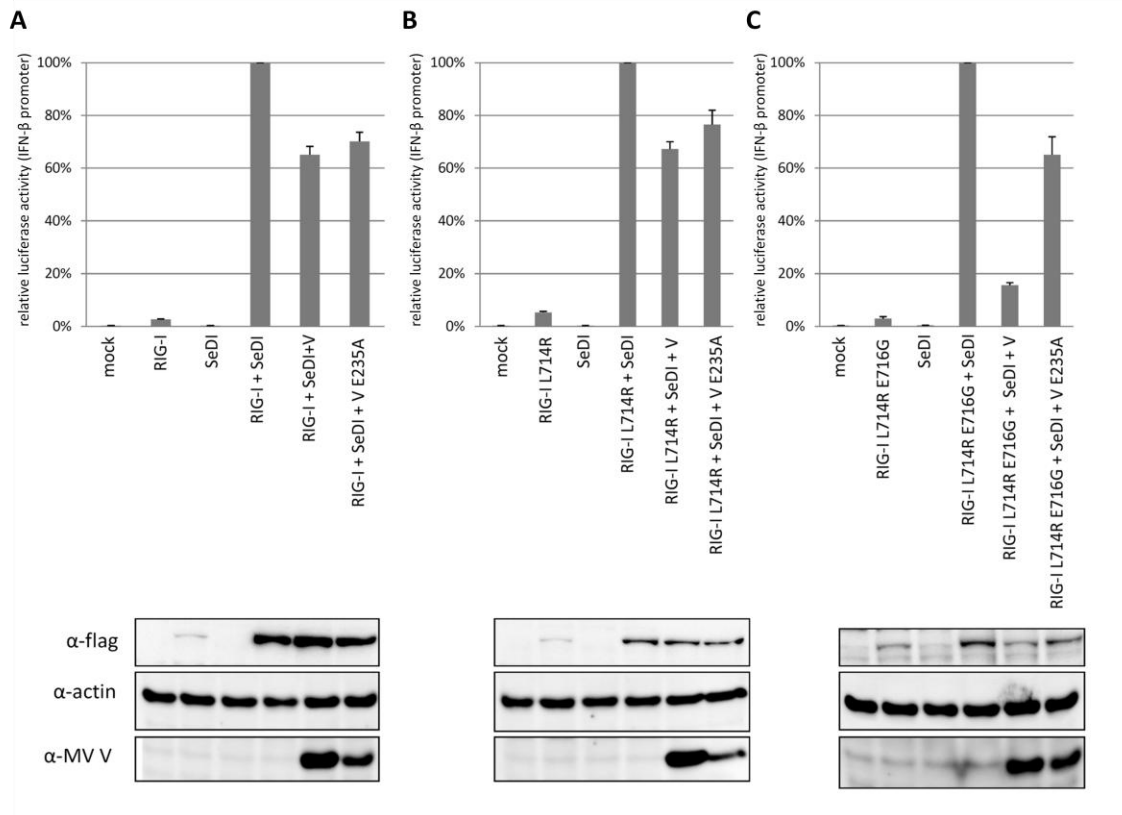
#### 4.3.1 Engineering of RIG-I inhibition by paramyxoviral V proteins

The question arises if the structural model presented in this thesis describes the recognition of MDA5 by V proteins sufficiently well to allow engineering not only of robust binding to V proteins into RIG-I but also V protein mediated inhibition of the RIG-I signaling activity in living cells. Figure 44 shows the activation of the interferon- $\beta$  promoter by stimulated wild type hsRIG-I and the structure-derived hsRIG-I mutants that mimic the MDA5 interface required for V protein binding. The reporter gene assays revealed that signaling mediated by wild type hsRIG-I is partially inhibited in presence of the wild type MV V protein as well as by the E235A<sup>MV</sup> mutant (see figure 44A). Wild type RIG-I was shown not to bind to the paramyxoviral V protein, neither *in vitro* (see figure 31B) nor in living cells (see figure 43). The observed partial inhibition of RIG-I signaling is probably caused by the multifunctionality of the V protein, which was shown to inhibit the induction of interferon- $\beta$  at several levels (109, 111-114) (see also section 1.5.2 and figure 6).

The gene reporter assays performed with the L714R<sup>hsRIG-I</sup> mutant and wild type MV V protein or E235A<sup>MV</sup> mutant demonstrated similar results as for the wild type hsRIG-I protein (see figure 44B). The L714R<sup>hsRIG-I</sup> mutant has been shown to bind the MV V protein in living cells (see figure 43), but not the PIV5 V protein in size exclusion chromatography experiments (see figure 31B with the equivalent L715G<sup>mmRIG-I</sup> mutant). However, the measured degree of L714R<sup>hsRIG-I</sup> signaling inhibition in presence of wild type MV V protein is comparable to wild type hsRIG-I and was not found to be attenuated in presence of the E235A<sup>MV</sup> mutant, indicating the L714R<sup>hsRIG-I</sup> single mutant is not sensitive to an inhibition induced by binding of the MV V protein. The E335A<sup>MV</sup> mutant was shown to abolish the induced V protein interaction of the engineered L714R<sup>hsRIG-I</sup> mutant in living cells (see figure 43) and serves therefore as a control in this experiment.

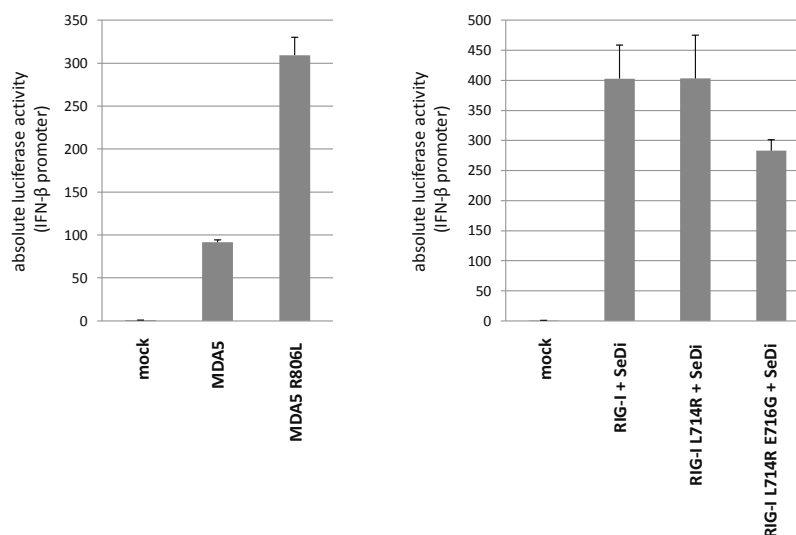
In comparison to the experiments performed with wild type hsRIG-I or with the L714R<sup>hsRIG-I</sup> mutant, the gene reporter assays examining the L714R E716G<sup>hsRIG-I</sup> mutant illustrate a distinct drop of its signaling activity in presence of wild type MV V protein (see figure 44C). This finding implies that the RIG-I double mutant is responsive to inhibition induced by MV V protein binding. The RIG-I double mutant has been shown to bind paramyxoviral V proteins in living cells (see figure 43) as well as in *in vitro* size exclusion chromatography experiments (see figure 31B with the equivalent L715R E717G<sup>mmRIG-I</sup> mutant). Unlike in the case of the L714R<sup>hsRIG-I</sup> mutant, the decrease of the signaling activity seen for the L714R E716G<sup>hsRIG-I</sup> mutant was found to be released in presence of the E235A<sup>MV</sup> mutant. This control experiment verifies that the observed drop in signaling activity can be indeed ascribed to recognition of the RIG-I double mutant by the MV V protein and demonstrates that the

structure-derived model of MDA5 inhibition is sufficient to engineer not only binding, but also inhibition by paramyxoviral V proteins into RIG-I.



**Fig. 44:** Signaling inhibition of wild type hsRIG-I or engineered RIG-I mutants (that mimic the MDA5 interface required for V protein binding) by MV V protein (unpublished data of Dr. Kerstin Monika Schuhmann and Prof. Karl-Klaus Conzelmann, Gene Center Munich). Shown are interferon- $\beta$  promoter dependent luciferase reporter gene assays performed in HEK-293T RIG-I knock out cells with a promoter sequence which is responsive to the transcription factors AP-1, NF- $\kappa$ B, IRF-3 and IRF-7 (153). Mean values ( $\pm$  SEM) were obtained from experiments performed in triplicate (N=3). Expression levels of transfected flag-tagged hsRIG-I (wild type or mutants) and MV V protein (wild type or mutant) plasmids were analyzed by western blot (lower panels). Flag-tagged hsRIG-I (wild type or mutants) was stimulated using Sendai virus defective-interfering genomes (SeDI) (160). **(A)** Reporter gene assay performed with wild type hsRIG-I and MV V WT or MV V E235A. Shown are relative values of the interferon- $\beta$  promoter activity related to the RIG-I WT value (100%). **(B)** Reporter gene assay performed with hsRIG-I L714R and MV V WT or MV V E235A. Shown are relative values of the interferon- $\beta$  promoter activity related to the RIG-I L714R value (100%). **(C)** Reporter gene assay performed with hsRIG-I L714R E716G and MV V WT or MV V E235A. Shown are relative values of the interferon- $\beta$  promoter activity related to the RIG-I L714R E716G value (100%).

The leucine and glutamate residue within the fifth  $\beta$ -strand of the RIG-I C-terminal RecA-like fold allows RIG-I to avoid recognition and inhibition by paramyxoviral V protein, but that alone does not explain the strict conservation of these two residues among RIG-I orthologs (see figure 30B). This raises the question of the actual function of the divergent residue conservation observed in MDA5 (R and G) and RIG-I (L and E). The comparison of the absolute activities of wild type MDA5 and the MDA5 mutant as well as wild type RIG-I and RIG-I mutants (see figure 45) measured in the described luciferase reporter gene assays (see figures 39 and 44) imply that these residues might play a role in the activity regulation of RIG-I like receptors. The MDA5 mutant, whose arginine is replaced by RIG-I's leucine, exhibits significantly higher activity than the unmodified wild type protein (see figure 45, right panel). Vice versa, the RIG-I double mutant, whose leucine and glutamate is replaced by MDA5's arginine and glycine, shows reduced activity compared to wild type RIG-I (see figure 45, left panel). It seems possible that these residues contribute to the communication between the different domains within the RNA clamp structure during the ATP hydrolysis cycle. Detailed elucidation of the course of action at the atomic level might give helpful insights into evolutionary constraints that underlie the distinct residue conservation within the SF2 ATPase domain or RLRs (see also section 1.3).



**Figure 45:** Left panel: Absolute activity of wild type hsMDA5 and the hsMDA5 R806L mutant measured in interferon- $\beta$  promoter dependent luciferase reporter gene assays performed in HEK-293T cells (see also figure 39). Right panel: Absolute activity of SeDi stimulated wild type hsRIG-I, the hsRIG-I L714R mutant and the hsRIG-I L714R E716G double mutant measured in Interferon- $\beta$  promoter dependent luciferase reporter gene assays performed in HEK-293T RIG-I knock out cells (see also figure 44). (Unpublished data of Dr. Kerstin Monika Schuhmann and Prof. Karl-Klaus Conzelmann, Gene Center Munich).

#### 4.3.2 Indirect inhibition by paramyxoviral V proteins via LGP2

Co-immunoprecipitation experiments have shown that MDA5 and LGP2 share the interactive site that mediates interference by paramyxoviral V proteins (135). Indeed, the LGP2 SF2 ATPase domain is able to form a stable, stoichiometric complex with the PIV5 V protein in size exclusion chromatography (see figure 30C). The arginine as well as the glycine residue identified to be critical for V protein binding by MDA5 are also conserved among LGP2 orthologs (see figure 30B). This implies similar interactions within the LGP2:V protein complex as seen for the ssMDA5:PIV5 V protein interface. It was further suggested that V protein binding to LGP2 mediates indirect inhibition of RIG-I by promoting the formation of inactive complexes between LGP2 and RIG-I (161). In addition, studies on LGP2 knockout mice let assume that LGP2 might function as a positive regulator of both MDA5 as well as RIG-I signaling (19) (see also section 1.4.3). Binding of paramyxoviral V proteins to LGP2 could affect this positive regulation of MDA5 and RIG-I during viral infection. However, to answer the question how the interaction of V proteins with LGP2 might contribute to an inhibition of MDA5 and RIG-I and to judge its impact on antiviral processes, requires detailed analysis of the actual function of LGP2 in RLR mediated signaling at the molecular level. In this context, it would be interesting to know if the LGP2 molecule itself is able to assemble into filamentous structures along viral RNA ligands and/or supports MDA5 filament formation *in vivo*.

#### 4.3.3 Viral pressure on the evolution of RLRs

Orthologs of MDA5, but not of RIG-I, exist in fish, indicating that MDA5 might have evolved before RIG-I (162). Furthermore, the absence of RIG-I in chicken is assumed to cause their increased susceptibility to influenza viruses (163). RIG-I has been shown to be the main PRR known for detecting influenza virus infection as well as infections by various members of the *Paramyxoviridae* family in mammalian cells (see also section 1.2) (14, 32-34). These findings let speculate that RIG-I might have evolved in higher organism to ensure efficient recognition and defense of virus species that are able to successfully defy sensing by MDA5, for instance through its inhibition by V proteins. It would be exciting to elucidate in detail how evolutionary pressure imposed by viruses directed the diversification of core functional domains of RLRs, which appears to be fundamental for a powerful viral surveillance system in vertebrates (see also sections 1.2 and 1.3).

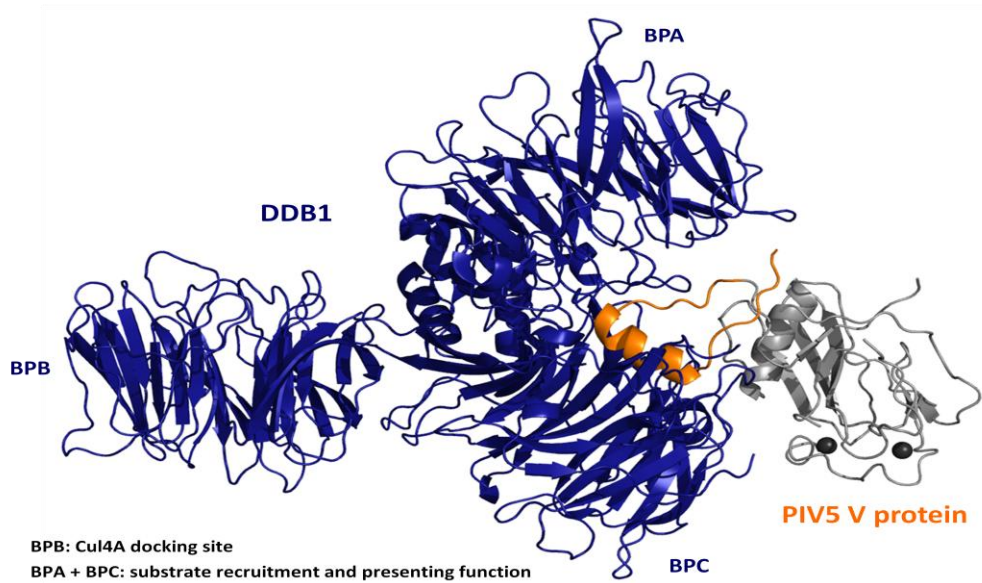
#### 4.4 The paramyxoviral V protein as a universal immune targeting tool

Numerous studies have shown that the V protein represents a pivotal element in immune evasion strategies of paramyxoviruses. This protein attacks not only the MDA5 signaling pathway by directly targeting MDA5 (see section 4.1), but also antagonizes common host factors of central immune pathways to severely disturb production as well as signaling of immunostimulating cytokines (see also section 1.5 and figure 6). Among these targets are the IKK $\epsilon$ /TBK1 kinase that activates the transcription factors IRF3 and IRF7 in TLR as well as in RLR signaling cascades (*114*), but also IRF3 and IRF7 themselves (*111*, *113*). Interference with the NF- $\kappa$ B family member p65 completes the inhibitory capacity of paramyxoviral V proteins for TLR and RLR signaling pathways at the transcriptional level (*109*). All these activities of the paramyxoviral V protein are directed against the induction of antiviral cytokine production, but the protein is also able to affect the paracrine and autocrine signaling system of cytokines either by direct inhibition of STAT transcription regulators or by mediating their proteasomal degradation (*104*, *115-129*).

The paramyxovirally induced degradation of STAT proteins requires binding of the V protein to DDB1 (*120*). DDB1 acts as a substrate recruiting adaptor protein within the DDB1-Cul4A ubiquitin ligase machinery that normally functions in the regulation of DNA repair, DNA replication and transcription (*104*, *125*). The V protein of mumps virus, PIV5 and hsPIV2 has been shown to capture DDB1 to exploit the ubiquitin ligase activity of the DDB1-Cul4A complex for STAT degradation (*118*, *119*, *122*). Structural determination of the PIV5 V protein bound to DDB1 revealed the molecular basis for this viral hijack (see figure 46) (*104*). In this complex, the PIV5 V protein was found to use an N-terminal  $\alpha$ -helical motif to interact with a  $\beta$ -propeller of DDB1 (BPC in figure 46). The PIV5 V protein inserts its helix into the binding pocket of the  $\beta$ -propeller that is usually responsible for the docking of DCAFs (DDB1-Cul4-associated factors) receptors, which mediate selective recruitment of specific substrates for ubiquitination by the DDB1-Cul4A E3 machinery (*164*). The C-terminal zinc binding motif (CTD) is not directly involved in the interaction with DDB1, but shielded by the NTD of the protein (*104*) (see also section 3.6). However, a tryptophan-rich motif within the CTD of paramyxoviral V protein has been assumed to provide association with various STAT proteins (*123*, *124*, *165*). In this way, the V protein could function as a bipartite adaptor that allows recruitment and presentation of STAT proteins to the DDB1-Cul4A ubiquitin ligase machinery to subvert its function to benefit viral replication.

The CTD of paramyxoviral V proteins is not only involved in binding of STAT factors and MDA5, but has also been shown to be sufficient for the interaction with the transcription factors IRF7 and p65 (*109*, *113*). Such findings imply that paramyxoviruses use their conserved CTD as a general binding module with various, but distinct destinations (e.g. MDA5 and LGP2 but not the third RLR family

member RIG-I, p65 but not the highly related p50 protein) to simultaneously counteract multiple major pathways of the innate immune system. Further structural analysis of the V protein in complex with different targets is required to reveal the impact of its versatility in detail and to deepen our knowledge of virus-host interactions and the evolutionary adaption of viral immune evasion strategies.



**Figure 46:** Crystal structure of the PIV5 V protein (shown in grey) in complex with DDB1 (shown in blue) with indicated functions of its  $\beta$ -propellers (BP) A, B and C (PDB code: 2B5L). The N-terminal  $\alpha$ -helical motif that the PIV5 V protein uses to bind to the BPC and thus to subvert the function of DDB1 as substrate recruiting adaptor protein within the DDB1-Cul4A ubiquitin ligase machinery is highlighted in orange. Adapted from (104, 125).

## 5. Summary

The innate immune system represents a universal first line defense of the host. It plays a crucial role in the early recognition of microbial pathogens and subsequent triggering of signaling cascades, which induce the defense reactions of the host cell during infection. To achieve this task, the innate immune system is armed with a large family of membrane-bound as well as cytoplasmic receptor proteins. These pattern recognition receptors (PRRs) mediate sensing of particular molecular signatures on pathogens called pathogen-associated molecular patterns (PAMPs). The specific interaction of the receptor molecule and its ligand allows the immune system to distinguish infectious nonself from noninfectious self.

The RIG-I like receptors (RLRs) RIG-I, MDA5 and LGP2 are key sentinels in the detection of viruses. They function as cytoplasmic PRRs by binding to viral RNAs during virus infection. The members share a modular domain organization composed of N-terminal tandem caspase activation and recruitment domains (CARDs) (which are absent in LGP2), a central superfamily 2 (SF2) ATP hydrolysis domain, and a C-terminal RNA binding domain (RD). Structural studies demonstrated that the RD and SF2 ATP hydrolysis domain of RIG-I as well as MDA5 act in concert to efficiently sense viral RNAs. They were found to form a RNA clamp structure, which encircles the viral dsRNA from all sites. This ring formation triggers the release or oligomerization of the CARDs and thereby transmission of the signal via a pathway that involves the interaction with the mitochondrial signaling adaptor protein MAVS. The initiated signal cascade leads finally to the activation of transcription factors such as IRFs and NF- $\kappa$ B, which control the production of type I interferons and other proinflammatory cytokines. Thus, RLRs play a central role in the induction of an antiviral state of the host cell.

Viruses have evolved numerous mechanisms to antagonize the host immune response. The family of paramyxoviruses produce the V protein, which specifically binds and inactivates MDA5, but not RIG-I. The highly conserved C-terminal domain (CTD) of these V proteins is known to fold into a unique zinc binding motif. Several studies showed that the integrity of this motif is required for binding and inhibition of the MDA5 ATP hydrolysis domain.

This thesis presents the crystal structure of the evolutionary conserved core of the MDA5:V protein complex and describes a first molecular mechanism of how a viral inhibitor protein targets an RLR signaling component. The structure of the Parainfluenza virus 5 (PIV5) V protein in complex with MDA5 revealed that paramyxoviral V proteins use their conserved CTD to integrate themselves into the ATP hydrolysis domain of MDA5. In detail, the V protein melts into the central  $\beta$ -sheet of the C-terminal RecA-like fold of MDA5 and displaces its canonical  $\beta$ -strands 1 and 6 by the two  $\beta$ -strands of the CTD. This partial unfolding of the MDA5 structure allows V proteins to trap MDA5 in an inactive state.



The wedging of the viral CTD into the C-terminal RecA-like fold was found to disrupt the conserved motif VI, which MDA5 shares with all other members of the SF2 helicase family and is known to be essential for ATP binding and hydrolysis. This explains how V proteins shut down the ATP hydrolysis activity of MDA5. Simultaneously, the insertion of the viral CTD causes a disruption of structural patterns within the SF2 ATP hydrolysis domain, which are required for the formation of the RNA clamp structure and hence for binding and proper sensing of viral RNA ligands by MDA5. As a consequence, the MDA5 protein was found to be unable to assemble into long filaments along viral dsRNA in the presence of the V protein. Cooperative filament formation of MDA5 on long dsRNA is a hallmark of this immune receptor and has to go hand in hand with ATP hydrolysis to ensure the assembly of a functional and active MDA5:MAVS signalosome, which finally triggers downstream pathways that limit viral replication. Targeting both, the ATPase as well as the RNA binding activity of MDA5, the paramyxoviral V protein represents a very potent inhibitor that can antagonize MDA5 mediated host immune response at the very first step of ligand recognition.

The structure of the MDA5:V protein core complex also explains why MDA5 is recognized by MDA5, but not the highly related RIG-I. The MDA5:V protein interface as well as binding studies revealed that the formation of a stable complex relies on a salt bridge between a highly conserved arginine residue within the C-terminal RecA-like fold of MDA5 and a highly conserved glutamate residue within the CTD of paramyxoviruses. RIG-I lacks the conserved arginine residue and has instead a conserved leucine residue at this position that provides no anchoring point for salt bridge formation with the viral glutamate residue. This difference in residue conservation allows RIG-I proteins to evade recognition and hence inhibition by paramyxoviral V proteins.

In conclusion, the structural and biochemical data presented in this thesis provide important mechanistic insights into the inhibition of MDA5 by paramyxoviral V proteins and enhances our knowledge of virus-host interactions and the evolutionary adaption of viral immune evasion strategies.



## 6. References

1. C. A. Janeway Jr, R. Medzhitov, Innate immune recognition. *Science Signalling* **20**, 197 (2002).
2. T. H. Mogensen, Pathogen recognition and inflammatory signaling in innate immune defenses. *Clinical Microbiology Reviews* **22**, 240 (2009).
3. R. Medzhitov, C. Janeway, Innate immunity: Minireview the virtues of a nonclonal system of recognition. *Cell* **91**, 295 (1997).
4. H. Kumar, T. Kawai, S. Akira, Pathogen recognition by the innate immune system. *International Reviews of Immunology* **30**, 16 (2011).
5. K. Newton, V. M. Dixit, Signaling in innate immunity and inflammation. *Cold Spring Harbor Perspectives in Biology* **4**, (2012).
6. Y. Kumagai, S. Akira, Identification and functions of pattern-recognition receptors. *Journal of Allergy and Clinical Immunology* **125**, 985 (2010).
7. E. M. Creagh, L. O'Neill, TLRs, NLRs and RLRs: a trinity of pathogen sensors that co-operate in innate immunity. *Trends in Immunology* **27**, 352 (2006).
8. A. Iwasaki, A Virological View of Innate Immune Recognition. *Annual Review of Microbiology* **66**, (2012).
9. T. Abe, *et al.*, STING Recognition of Cytoplasmic DNA Instigates Cellular Defense. *Molecular Cell*, (2013).
10. K. Malathi, B. Dong, M. Gale, R. H. Silverman, Small self-RNA generated by RNase L amplifies antiviral innate immunity. *Nature* **448**, 816 (2007).
11. P. Luthra, D. Sun, R. H. Silverman, B. He, Activation of IFN- $\beta$  expression by a viral mRNA through RNase L and MDA5. *Proceedings of the National Academy of Sciences* **108**, 2118 (2011).
12. M. Yoneyama *et al.*, The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nature immunology* **5**, 730 (2004).
13. D.-c. Kang, *et al.*, mda-5: An interferon-inducible putative RNA helicase with double-stranded RNA-dependent ATPase activity and melanoma growth-suppressive properties. *Proceedings of the National Academy of Sciences* **99**, 637 (2002).
14. Y.-M. Loo, M. Gale, Immune signaling by RIG-I-like receptors. *Immunity* **34**, 680 (2011).
15. A. M. Bruns, C. M. Horvath, Activation of RIG-I-like receptor signal transduction. *Critical Reviews in Biochemistry and Molecular Biology* **47**, 194 (2012).

16. Y. Cui, *et al.*, The Stat3/5 Locus Encodes Novel Endoplasmic Reticulum and Helicase-like Proteins That Are Preferentially Expressed in Normal and Neoplastic Mammary Tissue. *Genomics* **78**, 129 (2001).
17. D. Pollpeter, A. Komuro, G. N. Barber, C. M. Horvath, Impaired cellular responses to cytosolic DNA or infection with *Listeria monocytogenes* and vaccinia virus in the absence of the murine LGP2 protein. *PLoS ONE* **6**, e18842 (2011).
18. E. M. Y. Moresco, B. Beutler, LGP2: Positive about viral sensing. *Proceedings of the National Academy of Sciences* **107**, 1261 (2010).
19. T. Satoh *et al.*, LGP2 is a positive regulator of RIG-I–and MDA5-mediated antiviral responses. *Proceedings of the National Academy of Sciences* **107**, 1512 (2010).
20. A. Komuro, D. Bamming, C. M. Horvath, Negative regulation of cytoplasmic RNA-mediated antiviral signaling. *Cytokine* **43**, 350 (2008).
21. S. Rothenfusser *et al.*, The RNA helicase Lgp2 inhibits TLR-independent sensing of viral replication by retinoic acid-inducible gene-I. *The Journal of Immunology* **175**, 5260 (2005).
22. A. Murali *et al.*, Structure and function of LGP2, a DEX (D/H) helicase that regulates the innate immunity response. *Journal of Biological Chemistry* **283**, 15825 (2008).
23. D. Vitour, E. F. Meurs, Regulation of interferon production by RIG-I and LGP2: a lesson in self-control. *Science Signaling* **2007**, pe20 (2007).
24. O. Cordin, J. Banroques, N. K. Tanner, P. Linder, The DEAD-box protein family of RNA helicases. *Gene* **367**, 17 (2006).
25. A. Ranji, K. Boris-Lawrie, RNA helicases: emerging roles in viral replication and the host innate response. *RNA biology* **7**, 775 (2010).
26. D. Luo *et al.*, Structural insights into RNA recognition by RIG-I. *Cell* **147**, 409 (2011).
27. T. Nishino, K. Komori, D. Tsuchiya, Y. Ishino, K. Morikawa, Crystal Structure and Functional Implications of *Pyrococcus furiosus* Hef Helicase Domain Involved in Branched DNA Processing. *Structure* **13**, 143 (2005).
28. S. Cui *et al.*, The C-terminal regulatory domain is the RNA 5'-triphosphate sensor of RIG-I. *Molecular Cell* **29**, 169 (2008).
29. F. Jiang *et al.*, Structural basis of RNA recognition and activation by innate immune receptor RIG-I. *Nature* **479**, 423 (2011).
30. B. Wu *et al.*, Structural Basis for dsRNA Recognition, Filament Formation, and Antiviral Signal Activation by MDA5. *Cell*, (2012).

31. X. Li *et al.*, The RIG-I-like receptor LGP2 recognizes the termini of double-stranded RNA. *Journal of Biological Chemistry* **284**, 13881 (2009).
32. M. Schlee *et al.*, Recognition of 5' triphosphate by RIG-I helicase requires short blunt double-stranded RNA as contained in panhandle of negative-strand virus. *Immunity* **31**, 25 (2009).
33. Y.-M. Loo *et al.*, Distinct RIG-I and MDA5 signaling by RNA viruses in innate immunity. *Journal of Virology* **82**, 335 (2008).
34. H. Kato, K. Takahasi, T. Fujita, RIG-I-like receptors: cytoplasmic sensors for non-self RNA. *Immunological Reviews* **243**, 91 (2011).
35. L. Gitlin *et al.*, Essential role of mda-5 in type I IFN responses to polyriboinosinic: polyribocytidylic acid and encephalomyocarditis picornavirus. *Proceedings of the National Academy of Sciences* **103**, 8459 (2006).
36. A. Pichlmair *et al.*, Activation of MDA5 requires higher-order RNA structures generated during virus infection. *Journal of Virology* **83**, 10761 (2009).
37. J. P. Wang *et al.*, MDA5 and MAVS mediate type I interferon responses to coxsackie B virus. *Journal of Virology* **84**, 254 (2010).
38. S. A. McCartney *et al.*, MDA-5 recognition of a murine norovirus. *PLoS Pathogens* **4**, e1000108 (2008).
39. V. Hornung *et al.*, 5'-Triphosphate RNA is the ligand for RIG-I. *Science Signaling* **314**, 994 (2006).
40. A. Pichlmair *et al.*, RIG-I-mediated antiviral responses to single-stranded RNA bearing 5'-phosphates. *Science* **314**, 994 (2006).
41. J. B. Bowzard, P. Ranjan, S. Sambhara, RIG-I Goes Beyond Naked Recognition. *Cell Host & Microbe* **13**, 247 (2013).
42. S. J. Butcher, J. M. Grimes, E. V. Makeyev, D. H. Bamford, D. I. Stuart, A mechanism for initiating RNA-dependent RNA polymerization. *Nature* **410**, 235 (2001).
43. M. Weber *et al.*, Incoming RNA Virus Nucleocapsids Containing a 5'-Triphosphorylated Genome Activate RIG-I and Antiviral Signaling. *Cell Host & Microbe* **13**, 336 (2013).
44. J. Rehwinkel *et al.*, RIG-I detects viral genomic RNA during negative-strand RNA virus infection. *Cell* **140**, 397 (2010).
45. C. Lu *et al.*, The structural basis of 5' triphosphate double-stranded RNA recognition by RIG-I C-terminal domain. *Structure* **18**, 1032 (2010).

46. C. Lu, C. Ranjith-Kumar, L. Hao, C. C. Kao, P. Li, Crystal structure of RIG-I C-terminal domain bound to blunt-ended double-strand RNA without 5' triphosphate. *Nucleic Acids Research* **39**, 1565 (2011).
47. H. Kato *et al.*, Length-dependent recognition of double-stranded ribonucleic acids by retinoic acid-inducible gene-I and melanoma differentiation-associated gene 5. *The Journal of Experimental Medicine* **205**, 1601 (2008).
48. A. Peisley, S. Hur, Multi-level regulation of cellular recognition of viral dsRNA. *Cellular and Molecular Life Sciences*, 1 (2012).
49. R. Züst *et al.*, Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. *Nature Immunology* **12**, 137 (2011).
50. H. Kato *et al.*, Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses. *Nature* **441**, 101 (2006).
51. Q. Feng *et al.*, MDA5 Detects the Double-Stranded RNA Replicative Form in Picornavirus-Infected Cells. *Cell Reports*, (2012).
52. K. Triantafilou *et al.*, Visualisation of direct interaction of MDA5 and the dsRNA replicative intermediate form of positive strand RNA viruses. *Journal of Cell Science* **125**, 4761 (2012).
53. A. Peisley *et al.*, Cooperative assembly and dynamic disassembly of MDA5 filaments for viral dsRNA recognition. *Proceedings of the National Academy of Sciences* **108**, 21010 (2011).
54. I. C. Berke, Y. Modis, MDA5 cooperatively forms dimers and ATP-sensitive filaments upon binding double-stranded RNA. *The EMBO Journal* **31**, 1714 (2012).
55. F. Civril *et al.*, The RIG-I ATPase domain structure reveals insights into ATP-dependent antiviral signalling. *EMBO Reports* **12**, 1127 (2011).
56. X. Li *et al.*, Structural basis of double-stranded RNA recognition by the RIG-I like receptor MDA5. *Archives of Biochemistry and Biophysics*, (2009).
57. C. Motz *et al.*, Paramyxovirus V Proteins Disrupt the Fold of the RNA Sensor MDA5 to Inhibit Antiviral Signaling. *Science* **339**, 690 (2013).
58. Y. Wang *et al.*, Structural and functional insights into 5'-ppp RNA pattern recognition by the innate immune receptor RIG-I. *Nature Structural & Molecular Biology* **17**, 781 (2010).
59. K. Takahasi *et al.*, Solution Structures of Cytosolic RNA Sensor MDA5 and LGP2 C-terminal Domains: Identification of the RNA Recognition Loop in RIG-I-like Receptors. *Journal of Biological Chemistry* **284**, 17465 (2009).
60. E. Kowalinski *et al.*, Structural basis for the activation of innate immune pattern-recognition receptor RIG-I by viral RNA. *Cell* **147**, 423 (2011).

61. I. C. Berke, X. Yu, Y. Modis, E. H. Egelman, MDA5 assembles into a polar helical filament on dsRNA. *Proceedings of the National Academy of Sciences* **109**, 18437 (2012).
62. S. Myong *et al.*, Cytosolic Viral Sensor RIG-I Is a 5'-Triphosphate-Dependent Translocase on Double-Stranded RNA. *Science* **323**, 1070 (2009).
63. M. U. Gack *et al.*, TRIM25 RING-finger E3 ubiquitin ligase is essential for RIG-I-mediated antiviral activity. *Nature* **446**, 916 (2007).
64. M. U. Gack *et al.*, Roles of RIG-I N-terminal tandem CARD and splice variant in TRIM25-mediated antiviral signal transduction. *Proceedings of the National Academy of Sciences* **105**, 16743 (2008).
65. D. Gao *et al.*, REUL is a novel E3 ubiquitin ligase and stimulator of retinoic-acid-inducible gene-I. *PLoS ONE* **4**, e5760 (2009).
66. J. Maelfait, R. Beyaert, Emerging role of ubiquitination in antiviral RIG-I signaling. *Microbiology and Molecular Biology Reviews* **76**, 33 (2012).
67. W. Zeng *et al.*, Reconstitution of the RIG-I pathway reveals a signaling role of unanchored polyubiquitin chains in innate immunity. *Cell* **141**, 315 (2010).
68. X. Jiang *et al.*, Ubiquitin-induced oligomerization of the RNA sensors RIG-I and MDA5 activates antiviral innate immune response. *Immunity*, (2012).
69. T. Kawai *et al.*, IPS-1, an adaptor triggering RIG-I-and Mda5-mediated type I interferon induction. *Nature Immunology* **6**, 981 (2005).
70. E. Meylan *et al.*, Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus. *Nature* **437**, 1167 (2005).
71. R. B. Seth, L. Sun, C.-K. Ea, Z. J. Chen, Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF-kappaB and IRF 3. *Cell* **122**, 669 (2005).
72. L.-G. Xu *et al.*, VISA is an adapter protein required for virus-triggered IFN- $\beta$  signaling. *Molecular Cell* **19**, 727 (2005).
73. J. A. Potter, R. E. Randall, G. L. Taylor, Crystal structure of human IPS-1/MAVS/VISA/Cardif caspase activation recruitment domain. *BMC Structural Biology* **8**, 11 (2008).
74. F. Hou *et al.*, MAVS forms functional prion-like aggregates to activate and propagate antiviral innate immune response. *Cell* **146**, 448 (2011).
75. S.-C. Lin, Y.-C. Lo, H. Wu, Helical assembly in the MyD88-IRAK4-IRAK2 complex in TLR/IL-1R signalling. *Nature* **465**, 885 (2010).

76. M.-C. Michallet *et al.*, TRADD protein is an essential component of the RIG-like helicase antiviral pathway. *Immunity* **28**, 651 (2008).
77. S. Balachandran, E. Thomas, G. N. Barber, A FADD-dependent innate immune mechanism in mammalian cells. *Nature* **432**, 401 (2004).
78. H. Häcker *et al.*, Specificity in Toll-like receptor signalling through distinct effector functions of TRAF3 and TRAF6. *Nature* **439**, 204 (2005).
79. K. Takahashi *et al.*, Cutting edge: roles of caspase-8 and caspase-10 in innate immune responses to double-stranded RNA. *The Journal of Immunology* **176**, 4520 (2006).
80. F. Liu, Y. Xia, A. S. Parker, I. M. Verma, IKK biology. *Immunological Reviews* **246**, 239 (2012).
81. G. Ghosh, V. Y. F. Wang, D. B. Huang, A. Fusco, NF- $\kappa$ B regulation: lessons from structures. *Immunological Reviews* **246**, 36 (2012).
82. M. Hinz, S. Ç. Arslan, C. Scheidereit, It takes two to tango: I $\kappa$ Bs, the multifunctional partners of NF- $\kappa$ B. *Immunological Reviews* **246**, 59 (2012).
83. M. Hayden, A. West, S. Ghosh, NF- $\kappa$ B and the immune response. *Oncogene* **25**, 6758 (2006).
84. A. Dev, S. Iyer, B. Razani, G. Cheng, in *NF- $\kappa$ B in Health and Disease*. (Springer, 2011), pp. 115-143.
85. C. Schindler, D. E. Levy, T. Decker, JAK-STAT signaling: from interferons to cytokines. *Journal of Biological Chemistry* **282**, 20059 (2007).
86. D. E. Levy, I. J. Marié, J. E. Durbin, Induction and function of type I and III interferon in response to viral infection. *Current opinion in virology* **1**, 476 (2011).
87. D. A. Pippig *et al.*, The regulatory domain of the RIG-I family ATPase LGP2 senses double-stranded RNA. *Nucleic Acids Research* **37**, 2014 (2009).
88. T. Venkataraman *et al.*, Loss of DExD/H box RNA helicase LGP2 manifests disparate antiviral responses. *The Journal of Immunology* **178**, 6444 (2007).
89. N. K. Duggal, M. Emerman, Evolutionary conflicts between viruses and restriction factors shape immunity. *Nature Reviews Immunology* **12**, 687 (2012).
90. G. Le Negrate, Viral interference with innate immunity by preventing NF- $\kappa$ B activity. *Cellular Microbiology* **14**, 168 (2012).
91. D. W. Leung, C. F. Basler, G. K. Amarasinghe, Molecular mechanisms of viral inhibitors of RIG-I-like receptors. *Trends in Microbiology* **20**, 139 (2012).
92. T. Wolff *et al.*, Sabotage of antiviral signaling and effectors by influenza viruses. *Biological Chemistry* **389**, 1299 (2008).



93. A. G. Bowie, L. Unterholzner, Viral evasion and subversion of pattern-recognition receptor signalling. *Nature Reviews Immunology* **8**, 911 (2008).
94. M. Munir, TRIM proteins: another class of viral victims. *Science Signaling* **3**, jc2 (2010).
95. D. Young, L. Didcock, S. Goodbourn, R. Randall, Paramyxoviridae use distinct virus-specific mechanisms to circumvent the interferon response. *Virology* **269**, 383 (2000).
96. B. Gotoh, T. Komatsu, K. Takeuchi, J. Yokoo, Paramyxovirus strategies for evading the interferon response. *Reviews in Medical Virology* **12**, 337 (2002).
97. R. Chambers, T. Takimoto, Antagonism of innate immunity by paramyxovirus accessory proteins. *Viruses* **1**, 574 (2009).
98. R. W. Peluso, R. A. Lamb, P. W. Choppin, Polypeptide synthesis in simian virus 5-infected cells. *Journal of Virology* **23**, 177 (1977).
99. A. Kato, K. Kiyotani, Y. Sakai, T. Yoshida, Y. Nagai, The paramyxovirus, Sendai virus, V protein encodes a luxury function required for viral pathogenesis. *The EMBO Journal* **16**, 578 (1997).
100. S. M. Thomas, R. A. Lamb, R. G. Paterson, Two mRNAs that differ by two nontemplated nucleotides encode the amino coterminal proteins P and V of the paramyxovirus SV5. *Cell* **54**, 891 (1988).
101. S. Vidal, J. Curran, D. Kolakofsky, Editing of the Sendai virus P/C mRNA by G insertion occurs during mRNA synthesis via a virus-encoded activity. *Journal of Virology* **64**, 239 (1990).
102. R. G. Paterson, G. P. Leser, M. A. Shaughnessy, R. A. Lamb, The paramyxovirus SV5 V protein binds two atoms of zinc and is a structural component of virions. *Virology* **208**, 121 (1995).
103. P. Liston, D. J. Briedis, Measles virus V protein binds zinc. *Virology* **198**, 399 (1994).
104. T. Li, X. Chen, K. C. Garbutt, P. Zhou, N. Zheng, Structure of DDB1 in complex with a paramyxovirus V protein: viral hijack of a propeller cluster in ubiquitin ligase. *Cell* **124**, 105 (2006).
105. T. V. O'Halloran, Transition metals in control of gene expression. *Science* **261**, 715 (1993).
106. D. Karlin, S. Longhi, V. Receveur, B. Canard, The N-Terminal Domain of the Phosphoprotein of *Morbilliviruses* Belongs to the Natively Unfolded Class of Proteins. *Virology* **296**, 251 (2002).
107. B. He *et al.*, Recovery of paramyxovirus simian virus 5 with a V protein lacking the conserved cysteine-rich domain: the multifunctional V protein blocks both interferon- $\beta$  induction and interferon signaling. *Virology* **303**, 15 (2002).
108. E. Poole, B. He, R. A. Lamb, R. E. Randall, S. Goodbourn, The V proteins of simian virus 5 and other paramyxoviruses inhibit induction of interferon- $\beta$ . *Virology* **303**, 33 (2002).

109. K. M. Schuhmann, C. K. Pfaller, K.-K. Conzelmann, The measles virus V protein binds to p65 (RelA) to suppress NF- $\kappa$ B activity. *Journal of Virology* **85**, 3162 (2011).
110. T. Sakaguchi *et al.*, Analysis of interaction of Sendai virus V protein and melanoma differentiation-associated gene 5. *Microbiology and Immunology* **55**, 760 (2011).
111. T. Irie, K. Kiyotani, T. Igarashi, A. Yoshida, T. Sakaguchi, Inhibition of interferon regulatory factor 3 activation by paramyxovirus V protein. *Journal of Virology* **86**, 7136 (2012).
112. J. Ye, T. Maniatis, Negative regulation of interferon- $\beta$  gene expression during acute and persistent virus infections. *PLoS ONE* **6**, e20681 (2011).
113. C. K. Pfaller, K.-K. Conzelmann, Measles virus V protein is a decoy substrate for I $\kappa$ B kinase  $\alpha$  and prevents Toll-like receptor 7/9-mediated interferon induction. *Journal of Virology* **82**, 12365 (2008).
114. L. L. Lu, M. Puri, C. M. Horvath, G. C. Sen, Select Paramyxoviral V Proteins Inhibit IRF3 Activation by Acting as Alternative Substrates for Inhibitor of  $\kappa$ B Kinase  $\epsilon$  (IKKe)/TBK1. *Journal of Biological Chemistry* **283**, 14269 (2008).
115. J. J. Rodriguez, J.-P. Parisien, C. M. Horvath, Nipah virus V protein evades alpha and gamma interferons by preventing STAT1 and STAT2 activation and nuclear accumulation. *Journal of Virology* **76**, 11476 (2002).
116. J. J. Rodriguez, L.-F. Wang, C. M. Horvath, Hendra virus V protein inhibits interferon signaling by preventing STAT1 and STAT2 nuclear accumulation. *Journal of Virology* **77**, 11842 (2003).
117. H. Palosaari, J.-P. Parisien, J. J. Rodriguez, C. M. Ulane, C. M. Horvath, STAT protein interference and suppression of cytokine signal transduction by measles virus V protein. *Journal of Virology* **77**, 7635 (2003).
118. L. Didcock, D. Young, S. Goodbourn, R. Randall, The V protein of simian virus 5 inhibits interferon signalling by targeting STAT1 for proteasome-mediated degradation. *Journal of Virology* **73**, 9928 (1999).
119. J.-P. Parisien *et al.*, The V protein of human parainfluenza virus 2 antagonizes type I interferon responses by destabilizing signal transducer and activator of transcription 2. *Virology* **283**, 230 (2001).
120. J. Andrejeva, E. Poole, D. Young, S. Goodbourn, R. Randall, The p127 subunit (DDB1) of the UV-DNA damage repair binding protein is essential for the targeted degradation of STAT1 by the V protein of the paramyxovirus simian virus 5. *Journal of Virology* **76**, 11379 (2002).

121. N. Yokosawa, S.-i. Yokota, T. Kubota, N. Fujii, C-terminal region of STAT-1 $\alpha$  is not necessary for its ubiquitination and degradation caused by mumps virus V protein. *Journal of Virology* **76**, 12683 (2002).
122. C. M. Ulane, C. M. Horvath, Paramyxoviruses SV5 and HPIV2 assemble STAT protein ubiquitin ligase complexes from cellular components. *Virology* **304**, 160 (2002).
123. M. Nishio, D. Garcin, V. Simonet, D. Kolakofsky, The Carboxyl Segment of the Mumps Virus V Protein Associates with Stat Proteins *in vitro* via a Tryptophan-Rich Motif. *Virology* **300**, 92 (2002).
124. C. M. Ulane *et al.*, Composition and assembly of STAT-targeting ubiquitin ligase complexes: paramyxovirus V protein carboxyl terminus is an oligomerization domain. *Journal of Virology* **79**, 10180 (2005).
125. S. Angers *et al.*, Molecular architecture and assembly of the DDB1–CUL4A ubiquitin ligase machinery. *Nature* **443**, 590 (2006).
126. M. Barry, K. Fruh, Viral modulators of cullin RING ubiquitin ligases: culling the host defense. *Science Signaling* **2006**, pe21 (2006).
127. C. M. Horvath, Silencing STATs: lessons from paramyxovirus interferon evasion. *Cytokine & Growth Factor Reviews* **15**, 117 (2004).
128. C. M. Horvath, Weapons of STAT destruction. *European Journal of Biochemistry* **271**, 4621 (2004).
129. A. Ramachandran, C. M. Horvath, Paramyxovirus disruption of interferon signal transduction: STATus report. *Journal of Interferon & Cytokine Research* **29**, 531 (2009).
130. J. Andrejeva *et al.*, The V proteins of paramyxoviruses bind the IFN-inducible RNA helicase, mda-5, and inhibit its activation of the IFN- $\beta$  promoter. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 17264 (2004).
131. A. Ramachandran, C. M. Horvath, Dissociation of paramyxovirus interferon evasion activities: universal and virus-specific requirements for conserved V protein amino acids in MDA5 interference. *Journal of Virology* **84**, 11152 (2010).
132. H. Takaki *et al.*, Strain-to-strain difference of V protein of measles virus affects MDA5-mediated IFN- $\beta$ -inducing potential. *Molecular Immunology* **48**, 497 (2011).
133. A. Schaap-Nutt *et al.*, Identification of human parainfluenza virus type 2 (HPIV-2) V protein amino acid residues that reduce binding of V to MDA5 and attenuate HPIV-2 replication in nonhuman primates. *Journal of Virology* **85**, 4007 (2011).

134. K. Childs *et al.*, mda-5, but not RIG-I, is a common target for paramyxovirus V proteins. *Virology* **359**, 190 (2007).
135. J.-P. Parisien *et al.*, A shared interface mediates paramyxovirus interference with antiviral RNA helicases MDA5 and LGP2. *Journal of Virology* **83**, 7252 (2009).
136. K. Childs, J. Andrejeva, R. Randall, S. Goodbourn, Mechanism of mda-5 Inhibition by paramyxovirus V proteins. *Journal of Virology* **83**, 1465 (2009).
137. A. Kirchhofer, Structural and Functional Analysis of RIG-I Like Helicases / Modulating Spectral Properties of the Green Fluorescence Protein with Nanobodies, Dissertation, LMU, (2009).
138. P. V. Konarev, M. V. Petoukhov, V. V. Volkov, D. I. Svergun, ATSAS 2.1, a program package for small-angle scattering data analysis. *Journal of applied Crystallography* **39**, 277 (2006).
139. D. I. Svergun, M. V. Petoukhov, M. H. Koch, Determination of domain structure of proteins from X-ray solution scattering. *Biophysical Journal* **80**, 2946 (2001).
140. D. Franke, D. I. Svergun, DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *Journal of applied Crystallography* **42**, 342 (2009).
141. V. V. Volkov, D. I. Svergun, Uniqueness of ab initio shape determination in small-angle scattering. *Journal of applied Crystallography* **36**, 860 (2003).
142. W. Wriggers, P. Chacón, Using Situs for the registration of protein structures with low-resolution bead models from X-ray solution scattering. *Journal of applied Crystallography* **34**, 773 (2001).
143. Y. Kim *et al.*, Large-scale evaluation of protein reductive methylation for improving protein crystallization. *Nature Methods* **5**, 853 (2008).
144. M. S. Weiss, Introduction to anomalous scattering and its use in macromolecular crystallography, (2007) ([www.esrf.eu/events/conferences/embo2007/weiss\\_AnomScatt\\_2007.pdf](http://www.esrf.eu/events/conferences/embo2007/weiss_AnomScatt_2007.pdf))
145. G. Taylor, The phase problem, *Acta Crystallographica*, (2003)
146. Z. Dauter, M. Dauter, E. Dodson, Jolly SAD. *Acta Crystallographica Section D: Biological Crystallography* **58**, 494 (2002).
147. W. Kabsch, Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *Journal of applied Crystallography* **26**, 795 (1993).
148. C. Vonrhein, E. Blanc, P. Roversi, G. Bricogne, Automated structure solution with autoSHARP *Macromolecular Crystallography Protocols, Methods in Molecular Biology*, (2007)
149. K. Cowtan, The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallographica Section D: Biological Crystallography* **62**, 1002 (2006).

150. P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D: Biological Crystallography* **60**, 2126 (2004).
151. P. D. Adams *et al.*, PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography* **66**, 213 (2010).
152. F. H. Niesen, H. Berglund, M. Vedadi, The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature Protocols* **2**, 2212 (2007).
153. M. Yoneyama *et al.*, Autocrine amplification of type I interferon gene expression mediated by interferon stimulated gene factor 3 (ISGF3). *Journal of Biochemistry* **120**, 160 (1996).
154. H. Remaut *et al.*, Donor-strand exchange in chaperone-assisted pilus assembly proceeds through a concerted  $\beta$  strand displacement mechanism. *Molecular Cell* **22**, 831 (2006).
155. F. G. Sauer, J. S. Pinkner, G. Waksman, S. J. Hultgren, Chaperone priming of pilus subunits facilitates a topological transition that drives fiber formation. *Cell* **111**, 543 (2002).
156. F. Jacob-Dubuisson, R. Striker, S. J. Hultgren, Chaperone-assisted self-assembly of pili independent of cellular energy. *Journal of Biological Chemistry* **269**, 12447 (1994).
157. S. Ye, E. J. Goldsmith, Serpins and other covalent protease inhibitors. *Current Opinion in Structural Biology* **11**, 740 (2001).
158. P. R. Elliott, J.-P. Abrahams, D. A. Lomas, Wild-type  $\alpha$ 1-antitrypsin is in the canonical inhibitory conformation. *Journal of Molecular Biology* **275**, 419 (1998).
159. J. A. Huntington, R. J. Read, R. W. Carrell, Structure of a serpin–protease complex shows inhibition by deformation. *Nature* **407**, 923 (2000).
160. L. Strahle, D. Garcin, D. Kolakofsky, Sendai virus defective-interfering genomes and the activation of interferon-beta. *Virology* **351**, 101 (2006).
161. K. Childs, R. Randall, S. Goodbourn, Paramyxovirus V proteins interact with the RNA Helicase LGP2 to inhibit RIG-I-dependent interferon induction. *Journal of Virology* **86**, 3411 (2012).
162. D. Sarkar, R. DeSalle, P. B. Fisher, Evolution of MDA-5/RIG-I-dependent innate immunity: independent evolution by domain grafting. *Proceedings of the National Academy of Sciences* **105**, 17040 (2008).
163. M. R. Barber, J. R. Aldridge Jr, R. G. Webster, K. E. Magor, Association of RIG-I with innate immunity of ducks to influenza. *Proceedings of the National Academy of Sciences* **107**, 5913 (2010).
164. T. Li, E. I. Robert, P. C. van Breugel, M. Strubin, N. Zheng, A promiscuous  $\alpha$ -helical motif anchors viral hijackers and substrate receptors to the CUL4-DDB1 ubiquitin ligase machinery. *Nature Structural & Molecular Biology* **17**, 105 (2009).

165. M. Nishio *et al.*, Identification of paramyxovirus V protein residues essential for STAT protein degradation and promotion of virus replication. *Journal of Virology* **79**, 8591 (2005).

## 7. List of abbreviations

AIM2 - absent in melanoma 2  
 ASC - apoptosis-associated speck-like protein containing a CARD  
 ATP - adenosintriphosphate  
 Bis-Tris - bis(2-hydroxyethyl)amino-tris(hydroxymethyl)methane  
 bp - base pair  
 bt - bos taurus  
 CARD - caspase recruitment domain  
 cf - canis familiaris  
 CTD - C-terminal domain  
 Cul4A - cullin RING ubiquitin ligase 4A  
 DCAF - DDB1-Cul4-associated factors  
 DDB1 - DNA damage-binding protein 1  
 DTT - dithiothreitol  
 EMCV - encephalomyocarditis virus  
 EMSA - electrophoretic mobility shift assays  
 ER - endoplasmic reticulum  
 FADD - Fas-associated protein with death domain  
 gg - gallus gallus  
 HEPES - 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid  
 HIN - hemopoietic expression, interferon-inducibility, nuclear localization  
 IFI16 - interferon-inducible protein 16  
 I $\kappa$ B - inhibitor of  $\kappa$ B  
 IKK - inhibitor of  $\kappa$ B kinase  
 IL - interleukin  
 IPTG - isopropyl- $\beta$ -D-thiogalactopyranoside  
 IRF - interferon regulatory factors  
 ISG - interferon stimulated genes  
 LB - lysogeny broth  
 LGP2 - laboratory of genetics and physiology 2  
 LPS - lipopolysaccharides

## VII LIST OF ABBREVIATIONS

---

LRR - leucine rich repeats  
MAVS - Mitochondrial antiviral-signaling protein  
MDA5 - melanoma differentiation associated protein 5  
mam - macaca mulatta  
mm - mus musculus  
MV - measles virus  
Mw - molecular weight  
NF- $\kappa$ B - nuclear factor “kappa-light-chain-enhancer” of activated B-cells  
NLR - NOD-like receptors  
NLS - Nuclear localization signal  
NOD - nucleotide-binding and oligomerization domain  
NTD - N-terminal domain  
OAS - 2'-5' oligoadenylate synthase  
PAGE - polyacrylamide gel electrophoresis  
PAMP - pathogen-associated molecular patterns  
PCR - polymerase chain reaction  
*PEG* - polyethylene glycol  
PIV5 - parainfluenza virus 5  
Pol III - RNA polymerase III  
poly(I:C) - polyinosinic:polycytidylic acid  
PPR - pattern recognition receptor  
PYD - pyrin domain  
PYHIN - pyrin and HIN domain  
RCL - reactive center loop  
RdRp - RNA-dependent RNA polymerase  
RIG-I - retinoic acid inducible gene I  
RING - really interesting new gene  
RIP - receptor-interacting serine/threonine-protein kinase  
Riplet - RING finger protein leading to RIG-I activation  
RLR - RIG-I like receptor  
ROS - reactive oxygen species  
SAD - single wavelength anomalous dispersion  
SAXS - Small-angle X-ray scattering



## VII LIST OF ABBREVIATIONS

---

SDS - sodium dodecyl sulfate

SF2 - superfamily 2

ss - sus scrofa

STAT - signal transducers and activators of transcription

STING - stimulator of interferon genes

TBK1 - TANK-binding kinase 1

tg - taeniopygia guttata

TIR - Toll/interleukin- 1 receptor

TLC - thin layer chromatography

TLR - Toll-like receptors

TRADD - tumor necrosis factor receptor type-1 associated death domain

TRAF - tumor necrosis factor receptor associated factors

TRIM - tripartite motif

Tris - tris(hydroxymethyl)aminomethane

VDC - V protein dependent degradation complexes

## 8. Acknowledgments – Danksagung

An dieser Stelle möchte ich mich bei all jenen bedanken, deren Hilfe und Unterstützung zum Gelingen dieser Arbeit beigetragen haben.

Großer Dank gebührt zuallererst Prof. Dr. Karl-Peter Hopfner, in dessen Labor diese Arbeit entstanden ist. Er hat es stets verstanden, mir wissenschaftlichen Freiraum zu lassen, aber auch die Situationen erkannt, die Zuspruch und Unterstützung erforderten. Darüber hinaus ist er mir immer als gerechter, objektiver und diplomatischer Gruppenleiter begegnet, dem eine gute Arbeitsatmosphäre und das Miteinander seiner Mitarbeiter sehr am Herzen liegen.

Für die gute Zusammenarbeit und Kooperation während meiner Promotion möchte ich mich außerdem mit besonderem Nachdruck bei Kerstin Schuhmann und Prof. Dr. Karl-Klaus Conzelmann bedanken.

Allen ehemaligen und jetzigen Mitgliedern meiner Arbeitsgruppe danke ich für ihre Kollegialität, die dieses Labor zu einem sehr angenehmen Arbeitsplatz machen. Besonderer Dank gilt Gregor Witte, Matthew Bennett, Katja Lammens, Filiz Civril, Nadja Fenn, Maria Lucas und Kristina Lakomek für ihre vielen Ratschläge, das Zeitnehmen für Diskussionen, den Zuspruch und ihre beständige Hilfe im Laufe dieser Arbeit. Brigitte Kessler, Alexandra Schele, Manuela Moldt, Olga Fetscher sowie Stephanie Wolf danke ich für ihre vielfältige und unermessliche Hilfe in der Organisation des Laboralltags. Großen Dank möchte ich Manuela Moldt für ihre tatkräftige Unterstützung und für die außerordentlich gute Teamarbeit in diesem Projekt aussprechen.

Meiner Familie, ganz besonders meinen Eltern sowie meinem Freund danke ich für ihre Liebe, ihr Vertrauen und ihren starken Rückhalt. Euch verdanke ich so vieles, dass Worte dafür nicht ausreichen und Euch ist diese Arbeit von ganzem Herzen gewidmet.



## 9. Curriculum vitae

### Personal details

Name	Carina Beate Motz
Date of birth	September 16, 1983
Place of birth	Fulda, Germany
Nationality	German

### Education

1994 – 2003	Gymnasium Marienschule Fulda, Germany
-------------	---------------------------------------

### Academic Training

10/2003 – 09/2005	Studies of Biology (intermediate exam) at the University of Bremen, Germany
10/2005 – 06/2009	Studies of Biology (diploma) at the Eberhard Karls University Tübingen, Germany
09/2007 – 02/2008	Research assistant in the group of Prof. Dr. Dennis McCance Centre for Cancer Research & Cell Biology Queen's University Belfast, United Kingdom
10/2008 – 06/2009	Diploma thesis in the group of Dr. Elisa Izaurralde Max Planck Institute for Developmental Biology Tübingen, Germany
07/2009 – 02/2010	Research assistant in the group of Dr. Elisa Izaurralde Max Planck Institute for Developmental Biology Tübingen, Germany
since 04/2010	PhD thesis in the group of Prof. Dr. Karl-Peter Hopfner, Gene Center, Department of Biochemistry, Ludwigs-Maximilians-University Munich, Germany

